DOCUMENT RESUME

ED 418 987 TM 028 266

AUTHOR	Scheuren, Fritz; Li, Bonnie
TITLE	Intersurvey Consistency in NCES Private School Surveys for 1993-94. Working Paper Series.
INSTITUTION	Synectics for Management Decisions, Inc., Arlington, VA.
SPONS AGENCY	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-WP-96-27
PUB DATE	1996-11-00
NOTE	177p.
AVAILABLE FROM	U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Avenue, N.W., Room 400, Washington, DC 20208-5652.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	Classification; Elementary Secondary Education; *Estimation (Mathematics); *Least Squares Statistics; *National Surveys; Pilot Projects; *Private Schools; *Reliability
IDENTIFIERS	*Private School Questionnaire (NCES); *Schools and Staffing Survey (NCES)

ABSTRACT

This report provides empirical results of attempts to achieve consistency of estimates between two National Center for Education Statistics (NCES) surveys, the 1993-94 Private School Survey (PSS) and the Schools and Staffing Survey (SASS). Comparisons are made among statistical and computational procedures that may achieve the desired consistency between estimates. An overview describes the problem of consistency and why it matters. Sections 2 through 4 of this report explore the adjustment alternatives, with the main methods described in Section 2. These include the modified generalized least squares (GLS) approach evaluated in a pilot study and the variation of the GLS proposed by I. Olkin (1958). The third section presents results for nine independent applications for each category of the nine-category NCES typology for private schools. Section 4 presents conclusions, summarizing the work done, and making some recommendations. The new Olkin approach represents some improvements on the GLS approach previously tested, but results are not as encouraging as had been hoped. Section 5 presents references, and the algorithms used are illustrated in an appendix (Section 6). (Contains 36 tables, 40 figures, and 49 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 418 987

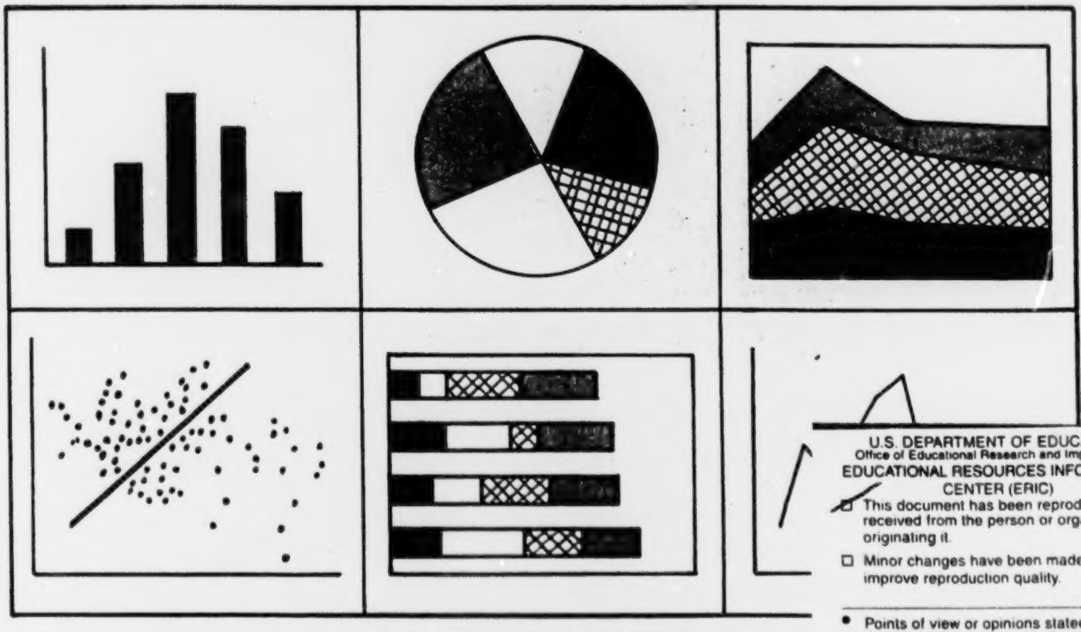
NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

Intersurvey Consistency in NCES Private School Surveys for 1993-94

Working Paper No. 96-27

November 1996



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

TMO28266

U.S. Department of Education
Office of Educational Research and Improvement

Intersurvey Consistency in NCES Private School Surveys for 1993-94

Working Paper No. 96-27 November 1996

Contact: Steven Kaufman
Surveys and Cooperative Systems Group
(202) 219-1337
e-mail: steve_kaufman@ed.gov

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Center for Education Statistics
Pascal D. Forgione, Jr.
Commissioner

Surveys and Cooperative Systems Group
Paul D. Planchon
Associate Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. *IntersurveyConsistency in NCES Private School Surveys for 1993-94.* Working Paper No. 96-27, by Fritz Scheuren and Bonnie Li. Steven Kaufman, project officer. Washington, D.C.: 1996.

November 1996

Foreword

Each year a large number of written documents are generated by NCES staff and individuals commissioned by NCES which provide preliminary analyses of survey results and address technical, methodological, and evaluation issues. Even though they are not formally published, these documents reflect a tremendous amount of unique expertise, knowledge, and experience.

The ***Working Paper Series*** was created in order to preserve the information contained in these documents and to promote the sharing of valuable work experience and knowledge. However, these documents were prepared under different formats and did not undergo vigorous NCES publication review and editing prior to their inclusion in the series. Consequently, we encourage users of the series to consult the individual authors for citations.

To receive information about submitting manuscripts or obtaining copies of the series, please contact Ruth R. Harris at (202) 219-1831 or U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Ave., N.W., Room 400, Washington, D.C. 20208-5654.

Susan Ahmed
Chief Mathematical Statistician
Statistical Standards and
Services Group

Samuel S. Peng
Director
Methodology, Training, and Customer
Service Program

Intersurvey Consistency in NCES
Private School Surveys for 1993-94

Fritz Scheuren
George Washington University

Bonnie Li
Synectics for Management Decisions, Inc.

November 1996

Synectics for Management Decisions, Inc.
3030 Clarendon Blvd., Suite 305
Arlington, VA 22201

Table of Contents

Foreword iii

Acknowledgments viii

1. **Overview and Background** 1

1.1 The Value of Studying Intersurvey Consistency 1
1.2 Private School Survey (PSS) Design 2
1.3 Private School Design in the Schools and Staffing Survey (SASS) 4
1.4 Selected Common Variables and their Definitions 5

2. **Initial Attempts at Achieving Intersurvey Consistency** 7

2.1 Modified Generalized Least Squares (GLS) Estimation 8
2.2 Olkin Variation of Basic GLS Approach 10
2.3 Discussion 11

3. **Results of GLS Applications by Type of School** 13

3.1 Catholic Parochial Typology 13
3.2 Catholic Diocesan Typology 28
3.3 Catholic Private Typology 42

3.4 Conservative Christian Typology 56
3.5 Other Religious Affiliated Typology 70
3.6 Other Religious Unaffiliated Typology 84

3.7 Nonsectarian Regular Typology 101
3.8 Nonsectarian Special Emphasis Typology 115
3.9 Nonsectarian Special Education Typology 129

4. **Summary and Recommendations** 143

4.1 Problem Restatement and Basic Approach 143
4.2 Summary of Empirical Results 144
4.3 Recommendations for Further Reweighting Research 148
4.4 Recommendations for Mass Imputation Research 150

5. **References** 154

6. **Appendix: Illustrations of Alternative Approaches** 158

Acknowledgments

The authors wish to thank Michael Chang for his assistance in formatting this document.

1. OVERVIEW AND BACKGROUND

This report provides empirical results of attempts to achieve consistency of estimates between two National Center of Education Statistics (NCES) surveys. These surveys are the 1993-94 Private School Survey (PSS) and the Private School Component of the 1993-94 Schools and Staffing Survey (SASS). Consistency was sought in the numbers of schools, teachers, and students from these two sources.

Comparisons have been made here among statistical and computational procedures that may serve to achieve the desired consistency between estimates. The complex nature of the PSS and SASS sample designs has also been considered, as well as any definitional differences which might exist between the surveys. In addition, potential benefits and the possibility of harm are addressed. The present work builds directly on an earlier pilot effort involving the 1991-92 PSS and the 1990-91 SASS (See Li and Scheuren 1995).

The goal of this overview section is to state the problem being addressed and why it may be important (in Subsection 1.1). An attempt is also made here to give sufficient background on the PSS so that the context and statistical issues are clear (in Subsection 1.2). For the same reasons, the design of the SASS private school component is discussed as well (in Subsection 1.3). To keep the treatment self-contained, definitions have been provided (Subsection 1.4).

In the following sections (sections 2 to 4) the adjustment alternatives are covered. The main methods being used are described in great detail (in Section 2). In Section 3, results from nine independent applications are given for each category of the nine NCES category typology for private schools (McMillen and Benson 1991); a concluding section, summarizing the work done and making some recommendations, ends the basic presentation (Section 4). References are included (Section 5) and an illustration is provided of the algorithms in an appendix (Section 6).

1.1 THE VALUE OF STUDYING INTERSURVEY CONSISTENCY

For the first time, in 1993-94, the private school component of the Schools and Staffing Survey (SASS) and the Private School Survey (PSS) were fielded in the same school year. Even though these two surveys measure some of the same variables, due to sampling and other errors, the results between the surveys did not agree.

As the PSS system is the basis for the SASS sampling frame, the PSS results, on the whole, are likely to be the more accurate. Under these circumstances, it makes sense to explore whether the introduction of 1993-94 PSS totals into the 1993-94 SASS might lead to improvements. Traditional post-stratification methods exist to employ auxiliary information at the estimation stage in surveys. These, however, cannot be applied to SASS without modification.

In particular, PSS and SASS both measure numbers of schools, numbers of teachers, and numbers of students. Conventional simple or raking ratio adjustment procedures could be used to adjust sample weights so that the SASS estimates agreed with PSS for each of the three totals separately (e.g., Oh and Scheuren, 1987). Such approaches do not work, though, if the weights are to be adjusted so that all three SASS estimates agree simultaneously. Other methods in the Generalized Least Squares (GLS) family, however, are available and, although new within an NCES framework, have proven to be of value elsewhere. Two of these are extensively studied in the present report and still other alternatives are discussed -- notably in the Recommendations (Section 4) and in the Appendix (Section 6).

1.2 PRIVATE SCHOOL SURVEY (PSS) DESIGN

The Private School Survey (PSS) is designed to collect data from all private schools in the 50 states and the District of Columbia. The survey is collected biennially by the U.S. Census Bureau for the National Center of Education Statistics (NCES).

Since 1983, NCES has used a dual frame approach for building its private school universe (U.S. Department of Education 1984). The dual frame consists of a list frame and an area frame.

The list building component (Broughman 1996) is the primary means for improving coverage of private schools. As Broughman states three major sources were employed to "build the list": commercial lists of schools, private school association lists, and state lists.

To identify schools that may have been overlooked in the list building component, an area frame was also included. The combination of the universe list and the additional schools identified in the area search comprised the cases included in the 1993-94 Private School Survey.

A more detailed description of each component of the dual frame is given below. The information which follows basically has been taken from Broughman (1996) and Broughman et al (1994).

1.2.1 List Frame.-- The starting point of the 1993-93 PSS list frame was the 1991-92 PSS. Additional steps were taken before fielding the 1993-94 PSS, though, to update and otherwise improve on this information.

To improve coverage of private schools in the list frame, before sending out the 1993-94 PSS, NCES requested and collected membership lists from 20 private school associations and denominations. NCES and Census also collected an updated list from the Quality Education Data or QED system plus lists of private schools from the 50 states, the District of Columbia, and Josten's, a company which sells school rings. Schools on private school association membership lists and the state lists were compared to the base list from the 1991-92 PSS. Any school on an association or denomination list, state list, the QED update list, or Josten's which did not match a school on the base list was added to the NCES private school universe list. As a result of these efforts, approximately 3,000 schools were added in 1993 to the NCES private school universe list (Jackson et al 1994).

1.2.2 Area Frame: First Stage.-- The area frame was designed to represent the private schools missing from the list frame. Additional sample schools were identified through an area search of randomly selected primary sampling units (PSUs).

The 1993-94 PSS area frame was designed to produce approximately a 50% overlap with the previous PSS. Consequently, the area frame consisted of two sets of sample PSU's: a subsample of the 1991-92 PSS area frame sample PSU's (overlap): and an independent sample of PSU's selected systematically with probabilities proportional to the square root of 1991 projected population.

The eight certainty PSUs in the 1991-92 PSS area frame remained in the 1993-94 SASS sample with certainty. For 1993-94 PSS, the schools in the 1991-92 certainty area frame PSUs were made a part of the list frame. Of the 60 PSUs in the 1991-92 PSS, there were 58 PSUs that had been in 1990-91 PSS for the first time and not previously been overlapped; these were selected again for the 1993-94 PSS, thus becoming the 1993-94 overlap sample of PSUs.

An additional 58 PSUs were selected independently. The United States was divided up into 2054 primary sampling units (PSUs). Each PSU consisted of a single county, independent city or cluster of geographically contiguous areas defined so that each PSU had a minimum population of 20,000 according to population projections for 1988, when the PSUs were first formed. To avoid having PSUs covering too large a geographic area, in a few cases some PSUs had less.

The strata were defined the same way as in the 1991-92 PSS area frame design: a) Census region (4 levels -- Northeast, Midwest, South, and West), b) metro/nonmetro status (2 levels) and c) whether the PSUs percent private school enrollment exceeded the median percent private enrollment of the other PSUs in the Census region/metro status strata (2 levels - using 1980 Census data).

A minimum of two PSUs were allocated to each of the 16 strata (32 PSUs). Also 26 additional PSUs were allocated to the 16 strata to more nearly approximate a uniform sampling fraction of PSUs from each stratum.

The nonoverlap PSUs were selected as a systematic sample with probability proportionate to the square root of the 1991 projected PSU population. A total of 123 distinct PSUs were in sample since one PSU was selected for both sets of samples. Its weight was adjusted to appropriately reflect the duplication.

1.2.3 Area Sample Frame: Within PSU Construction. --Within each of the 123 PSUs, the Census Bureau attempted to find all eligible private schools (i.e., nonpublic schools providing the following: instruction for any grades 1 -12, instruction not provided exclusively in the home, and a normal school day at least 4 hours long). An area canvas was not attempted. However, regional field staff created the frame using such sources as: yellow pages, non-Roman Catholic religious institutions, local education agencies, Chamber of Commerce, and local government offices. Roman Catholic religious institutions were not contacted because the National Catholic Education Association provides a very complete list of parochial Catholic schools. Once these lists of schools were constructed, they were matched with the updated 1993-94 list frame school file. Schools that did not match the list were considered part of the area frame.

For 1993-94, a total of 355 additional schools were found in the area sample; of these, 158 were found in PSU's not selected with certainty (153 after removing duplications). They were all included in sample as part of the area frame. The remaining 197 schools were in counties selected with certainty; and, hence, could be added to the list frame before the selection of the school sample.

1.2.4 Combined List and Area Samples.-- Data collection for the 1993-94 PSS was completed in March 1994. The final response rate was 91.8 percent . Of the 28,229 schools selected in the combined sample, some 3,741 cases were considered out-of-scope. The final weighted total of in-scope schools was 26,067 -- with 24,067 weighted schools coming from the list frame and 2,026 weighted cases from the area frame (after unduplication).

1.3 PRIVATE SCHOOL DESIGN IN THE SCHOOLS AND STAFFING SURVEY (SASS)

For the 1993-94 SASS, the private school portion was also selected using a dual frame approach -- analogous to that for the 1993-94 PSS. The 1993-94 SASS list frame can be considered simply a subsample of the 1993-94 PSS list cases. For the area frame, because of operational timing issues, this was not possible. A detailed description has been provided below, taken basically from Abramson et al 1996(See also Kaufman and Huang 1993).

1.3.1 List Frame.-- The 1993-94 SASS list frame used for private schools was the 1991-92 Private School Survey (PSS) list frame before any updating with additional (association) lists. Before sampling, duplicate schools were excluded from the frame. Schools that only teach prekindergarten, kindergarten or adult education were also removed. After sampling additional duplicates were discovered and eliminated as well.

1.3.2 Area Frame.-- The area frame sample consisted of two sets of sample PSUs: (1) a subsample of the area frame PSU's selected from the 1991-92 PSS (overlap); and (2) a sample of PSUs selected independently from the 1991-92 PSS area frame PSUs described in Section 1.2 above. By maintaining a fifty percent overlap of PSUs, the reliability of estimates of change was maintained at a reasonable level, while reducing respondent burden.

1.3.3 Combined List and Area Samples.-- Data collection for the 1993-94 SASS was completed in June of 1994. Of the 3,315 schools selected in the combined sample, some 241 cases were considered out-of-scope, 2,585 schools were respondents and 489 schools were not respondents. The final weighted SASS total of in-scope schools was 26,093 -- with 24,767 weighted schools coming from the list frame and 1,326 cases from the SASS area frame (after unduplication).

It may be worth commenting that the list portion of the PSS, as a universe count, is definitely to be relied upon in any attempts at achieving intersurvey consistency. The area portions of the PSS and SASS are both samples; and, hence, each has inherent variability. Adjusting the smaller SASS area sample to the larger PSS area sample will help but adjusting both to some combination of the two might be preferable to just relying on the PSS alone.

In this report, however, the PSS totals were taken as fixed and known with certainty. In later SASS applications, other approaches will be recommended, including separating the coverage adjustment in SASS from the survey itself. This point will be returned to later (See Section 4 and also Kaufman and Scheuren 1996).

1.4 SELECTED COMMON VARIABLES AND THEIR DEFINITIONS

Listed below are definitions of the key variables used in this report. These have been taken from several NCES reports -- notably Broughman (1996) and Broughman et al.(1994) plus McMillen and Benson (1991). The typology classification is listed first. Definitions for school, teacher, and student follow.

1.4.1 Typology.-- For the private school population, a typology exists which starts with the categorization (Catholic, Other Religious, and Nonsectarian), and further subdivides each group into three additional groups:

Catholic
- o Parochial
- o Diocesan
- o Private

Other Religious
- o Affiliated with a conservative Christian school association
- o Affiliated with national denomination or other religious school association
- o Unaffiliated

Nonsectarian
- o Regular programs
- o Special emphasis
- o Special education

Among Catholic schools, the governance categories (Parochial, Diocesan, Private) are strongly tied to differences in curriculum, student population characteristics, program emphasis, and sources of revenue.

In the case of Other Religious schools, recent work documents major differences in decisionmaking, educational goals, revenue, and enrollment trends between denomination schools (e.g., Lutheran, Jewish, Seventh-day Adventist) and those non-denominational schools affiliated with a Conservative Christian school association (e.g., Accelerated Christian Education, American Association of Christian Schools, Association of Christian Schools International, Oral Roberts Educational Fellowship). Schools in this latter type are commonly known as evangelical or fundamental, and are not tied to a denomination per se, but rather are governed by a single church, a foundation, or a local society. A third Other Religious category, Unaffiliated, is included to capture those religious schools which affiliate with neither a national denomination nor with a conservative Christian school association.

The three nonsectarian school categories are determined not by governance but by program emphasis. This classification disentangles private schools offering a conventional academic program (Regular) from those which either serve special needs children (Special Education) or provide a program with a Special Emphasis (e.g., Arts, Vocational, Alternative).

1.4.2 Private School. -- A school is an institution for instruction which has (1) a minimum school day of four hours per day, (2) a minimum of 160 days per year, (3) at least a first grade or higher, and (4) one or more teachers. A private school is an institution which provides instruction for any of grades 1-12, has one or more teachers to give instruction, is not administered by a public agency, and is not operated in a private home.

1.4.3 Teacher. -- In general, any full-time or part-time teacher whose school reported that his or her primary assignment was teaching in any of grades K-12. In other words, a headcount has been employed. See table 13 (page 18) of Broughman (1996) for the comparable concept and further PSS tabular detail.

1.4.4 Student.-- Individuals identified in the PSS or SASS as enrolled in a private school for instruction in a pre-kindergarten, kindergarten, grades 1 to 12, ungraded or post-secondary class. In defining whether or not a school was eligible or not to be in PSS/SASS, it had, however, to have at least one grade in 1-12.

2. INITIAL ATTEMPTS AT ACHIEVING INTERSURVEY CONSISTENCY

For NCES Private School Surveys alternatives do exist which permit simultaneous consistency or near consistency in totals for schools, teachers, and students. In particular, the Generalized Least Squares (GLS) techniques advocated by Deville and Särndal (1992) can be used, as in Imbens and Hellerstein (1993). While the asymptotic properties of GLS and GLS-like estimators are attractive, their finite sampling properties are not necessarily desirable. Possible operational concerns with GLS procedures include:

(1) Some of the resulting weights may be less than one (and may even be negative).

(2) The procedure may be difficult to implement (when excessively small weights exist).

(3) Also, the effect on estimates not directly adjusted is unknown (and could be harmful).

The initial work on GLS estimators might be said to date at least back to Deming and Stephan (1942). A near complete set of references through most of the 1970s can be found in Oh and Scheuren (1978b). Among the most important of these is that by Ireland and Kullback (1968) which gives the first convergence proof for the original Deming-Stephan algorithm.

Major recent papers include, Bankier (1990); Brewer (1995), Deville and Särndal (1992); Deville, Särndal, and Sautory (1993); Fuller et al (1994), Imbens and Hellerstein (1993), Jayasuriya and Valliant (1995), Kott (1996), Little(1991), plus Little and Wu(1991). The recent book, entitled *Model Assisted Survey Sampling*, by Särndal, Swensson, and Wretman (1992) is an important source too.

Except for Oh and Scheuren (1978a) and Imbens and Hellerstein (1993), the GLS applications covered have been univariate in nature. Now, as already mentioned, in the SASS setting the problem is inherently three-dimensional: Schools, Teachers, and Students -- each of which needs to agree with an independent PSS total.

In the main body of this report two alternatives will be covered. Both are variants of the approach in Imbens and Hellerstein (1993), as suggested independently by Burton(1989):

-- The basic modified GLS approach is described first (section 2.1), as originally proposed and employed in NCES Working Paper No. 95-16 (Li and Scheuren 1995).

-- Problems uncovered with the basic GLS approach lead to an alternative, which is a variant (see section 2.2)of an idea by Olkin (1958).

-- This section concludes with a discussion of possible evaluation criteria, leading naturally into the applications to follow.

In an appendix at the end of the report there is a completely worked illustration of the computations. Also found there is still another GLS variant that was considered but had to be discarded.

2.1 MODIFIED GENERALIZED LEAST SQUARES (GLS) ESTIMATION

To discuss the basic algorithm employed in Generalized Least Squares, it is necessary to define some notation; in particular --

w_i	is the original SASS Private School base weight for the ith SASS observation, $i=1,...,n$.
t_i	is the SASS total of teachers for ith SASS observation, $i=1,...,n$.
s_i	is the SASS total of the students for the ith SASS observation, $i=1,...,n$.
N	is the total estimated number of schools, as given by PSS.
T	is the total estimated number of teachers, as given by PSS.
S	is the estimated total number of schools, as given by PSS.

In reweighting SASS three constraints are imposed on the new weights u_i,

$$\sum u_i = N$$

$$\sum u_i t_i = T$$

$$\sum u_i s_i = S$$

For our application the new weights u_i, subject to these constraints, are to be chosen, as in Burton (1989), to minimize a loss function which can be written as the sum of squares:

$$\sum (u_i - w_i)^2$$

This is perhaps the simplest and most straightforward loss function that might be chosen. Motivating it here is outside our present scope, except to say that the sensitivity of the final results to the loss function chosen (e.g.,Deville and Särndal, 1992, Deville et al., 1993) seems not to be too great (but this is, in part, an application issue and will be among the areas for future study, as set forth at the end of this report).

Now the usual Lagrange multiplier formulation of this problem yields after some algebra that the new weights are of the form:

$$u_i = w_i + \lambda_1 + \lambda_2 t_i + \lambda_3 s_i ,$$

where the λ's are obtained from the matrix expression

$$\underline{d} = M\underline{\lambda}$$

with the vector $\underline{d}$ consisting of three elements, each a difference between the corresponding PSS and SASS totals for schools (first component), teachers (second component), and students (third component); in particular

$$\begin{matrix} N - \sum w_i \\ T - \sum w_i t_i \\ S - \sum w_i s_i \end{matrix}$$

where the summations are over the SASS sample observations and the quantities: N, T, and S are known PSS totals for schools (N), teachers (T), and students (S) respectively.

The matrix M is given by:

$$\begin{matrix} n & \sum t_i & \sum s_i \\ \sum t_i & \sum t^2_i & \sum t_i s_i \\ \sum s_i & \sum t_i s_i & \sum s^2_i \end{matrix}$$

and $\underline{\lambda}$ is the vector of unknown GLS adjustment factors obtained from:

$$\lambda = M^{-1}\underline{d}$$

(Notice that the M matrix is based solely on the unweighted sample relationships among schools, teachers and students. This is not an essential feature of our approach; and, indeed, had we used another loss function, a weighted version of the M matrix could have been used, as is discussed later in this report.)

2.2 OLKIN VARIATION OF BASIC GLS APPROACH

Based on concerns raised in our pilot application of GLS, it seemed worthwhile to see if a reweighting step could be introduced before the GLS algorithm was employed. An old idea of Olkin(1958) formed our starting point.

Assume we have a total τ, say, of student enrollment in the current application. Suppose further, as is the case here, that this is to be estimated from a sample. Olkin proposed a multivariate ratio estimator for τ which, in our case, can be written as

$$Y = a_1 R_1 w_{.} + a_2 R_2 t_{.} + a_3 R_3 s_{.}$$

where the a_i are positive and add to 1, the x_i are sample totals and the R_i are conventional ratios estimated from the sample of τ and X_i of the form

$$R_1 = S/N$$

$$R_2 = S/T$$

$$R_3 = S/S$$

where

the lower case Roman letters w., t., and s. are the sample (SASS) estimates and

the upper case Roman letters N, T, and S are the target (PSS) values to be attained.

In the present report, the a_i are simply chosen to be equal to one-third; however, a more natural approach would be to select them so as to minimize the variance of Y. Given the complex sample design of SASS, though, this has been left for the future.

In principle, an Olkin adjustment to the original weights could be produced within whatever domain is desired; then in order to determine the "new" weight for that domain, all the cases would be adjusted such that they would have new weights

$$u_i = rw_i$$

where the overall ratio r is obtained by taking Y and dividing it by the corresponding estimate obtained from the original sample.

The intuition is that if the Olkin estimation is first carried out for small (appropriate) subdomains, then there would be a direct benefit from this step in those subdomains. The "r" adjustment has the effect of making a weighted convex combination of the d_i's equal to zero. Intuitively, this was expected to reduce the number of negative weights; and, when done separately within subclasses, to achieve some of the usual benefits of post-stratification. Finally, because the overall PSS/SASS differences shrink appreciably, the Olkin adjustment would be expected to minimize any harm that GLS might do.

In the section which follows, we have tested our greatly simplified Olkin-like approach to GLS by applying it over suitable subdomains by school size (leaving for later, as already mentioned, a way to choose the a_i so as to minimize the variance of the estimator).

2.3 DISCUSSION

So far the GLS algorithms have been discussed as if the issues were simply computational. In point of fact, the real challenges arising in any SASS implementation require statistical judgments. Among these are:

- Deciding on the level of SASS at which the constraints are to be imposed. For example, from a subject-matter perspective, it seems appropriate to do GLS estimation separately within the nine private school typologies (as done for this report). For some of the larger typologies, moreover, maybe even finer groupings might be attempted (say, school level or urbanicity). At what point will the potential benefits of a GLS adjustment outweigh the harm? (See Subsection 3.6 for an example where the GLS was applied below the typology level.)

- An issue with the Olkin-like version of GLS is choosing suitable subdomains for the initial ratio adjustment to be employed before the GLS reweighting. Here we have used domains defined by school size within typologies. Had it been possible domains by type of locality could also have been tried. Indeed, a two-dimensional Olkin-like adjustment might have been done using a raking version of our proposal (a point that will be returned too in the concluding section).

- Avoiding GLS weights u_i that are negative or too small (i.e., given that each SASS observation always represents at least itself, a natural requirement to impose is that $u_i \geq 1$ for all i). This concern is particularly troublesome because of the seemingly ad hoc flavor of what may be needed to get acceptable weights (however, see Huang, 1978). Since in SASS many of the largest schools have weights near one this problem led us to propose a partial imputation strategy. In particular, for the largest SASS schools, GLS reweighting will not be carried out; instead, a direct use of the PSS cases is envisioned where, through statistical matching of SASS with PSS, the SASS data will be imputed onto one or more of the PSS observations. (See Scheuren 1996 and Section 4 of this report for more discussion).

Clearly, concepts like "benefit" and "harm" are not uniquely defined. In the formulation here, "benefits" will include not only intersurvey consistency between SASS and PSS but also usual criteria like reductions in the mean square error of estimates not constrained directly. The concept of "harm" is somewhat more elusive still. Among the factors to consider are obvious ones such as --

- How difficult (expensive) is the method to implement, including to explain?

- How sensitive are unconstrained estimates to seemingly small but arbitrary decisions in the way the method is applied?

A measure of "harm" that grows directly out of GLS is to look at what is happening to the variances of the weights as successive constraints are applied. A variant of this is to examine the ratio of the sums of the squared weights (where adjusted is divided by original),

$$\sum u_i^2 / \sum w_i^2 \quad .$$

The intuitive notion here is that the larger this ratio the greater the possible harm to a statistic not correlated with the quantities being constrained. This is the approach taken in the Appendix. The range of the weights is also another indicator of harm (Li and Scheuren 1995). Mainly, though, we will be using regression measures to study what impacts the adjustments have in the variability of the weights (See Section 3).

To look at the mean square error of the GLS estimators obtained in SASS, a direct comparison will also be made (as in Section 3 below) to selected comparable PSS quantities not directly used in the GLS process. This so-called independent assessment will involve data by urbanicity and school size -- items, in part at least, not used in the GLS adjustment.

3. RESULTS OF GLS APPLICATIONS BY TYPE OF SCHOOL

For the nine major types of private schools, there has been an attempt to employ the modified GLS algorithm discussed in Section 2. These applications were done separately and have been reported on as such here.

The approach taken in all instances is the same. The presentation begins with an overall description for a typology of the PSS and original SASS totals for schools, teachers, and students; next there is an in-depth look at the relationship between teacher and student totals in the two sources. This is followed by a documentation of how the weight adjustment factors, the λ's, were derived (plus what they mean for the particular typology). The actual operating characteristics of the resulting weights are then extensively commented on. An independent assessment (by community type and school size) of the adjustments on variables not directly impacted comes next. Comparisons to the 1990-91 pilot work have been deferred to Section 4.

Each typology can be read as a case study, standing alone. Familiarity with scatterplot matrices (e.g., Cleveland, 1993) is assumed; beyond that, there are no special analytic tools used that are not either well-known or explained as they are taken up.

Comparisons across typologies are left to the Summary and Recommendations Section (Section 4). It is fair to say, though, that on the whole the Olkin GLS calculations were reasonably successful. Our expectations were both that they would lead to improvements in SASS mean square error and that operational difficulties would be lessened. The partial results obtained so far bear this out.

3.1 CATHOLIC PAROCHIAL TYPOLOGY

The Catholic Parochial typology represents the largest single type of private school. For example, in the 1993-94 Private School Survey, there were an estimated 5,127 Catholic Parochial schools or about 20% of the private school total for that year.

In table 1.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the school totals are already very close (within 0.7%); but that SASS has many fewer teachers and students than are shown in PSS (3.2% and 1.7% less respectively).

To set the stage for the calculations that follow, it might be worth looking at figure 1.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black □'s. Overlaying the PSS data and shown as gray Δ's are the data from the 407 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS and along the same axis. In fact, the slope of the student/teacher relationship is 17.5 for PSS and 17.2 for SASS -- virtually indistinguishable. While not directly comparable because of differences in sample designs, the R^2

Table 1.1 -- Catholic Parochial: Weighted schools totals before excluding outliers (Based on 4,964 PSS and 407 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	5,127	5,092	35
Teachers	79,736	77,168	2,568
Students	1,409,828	1,385,587	24,241

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

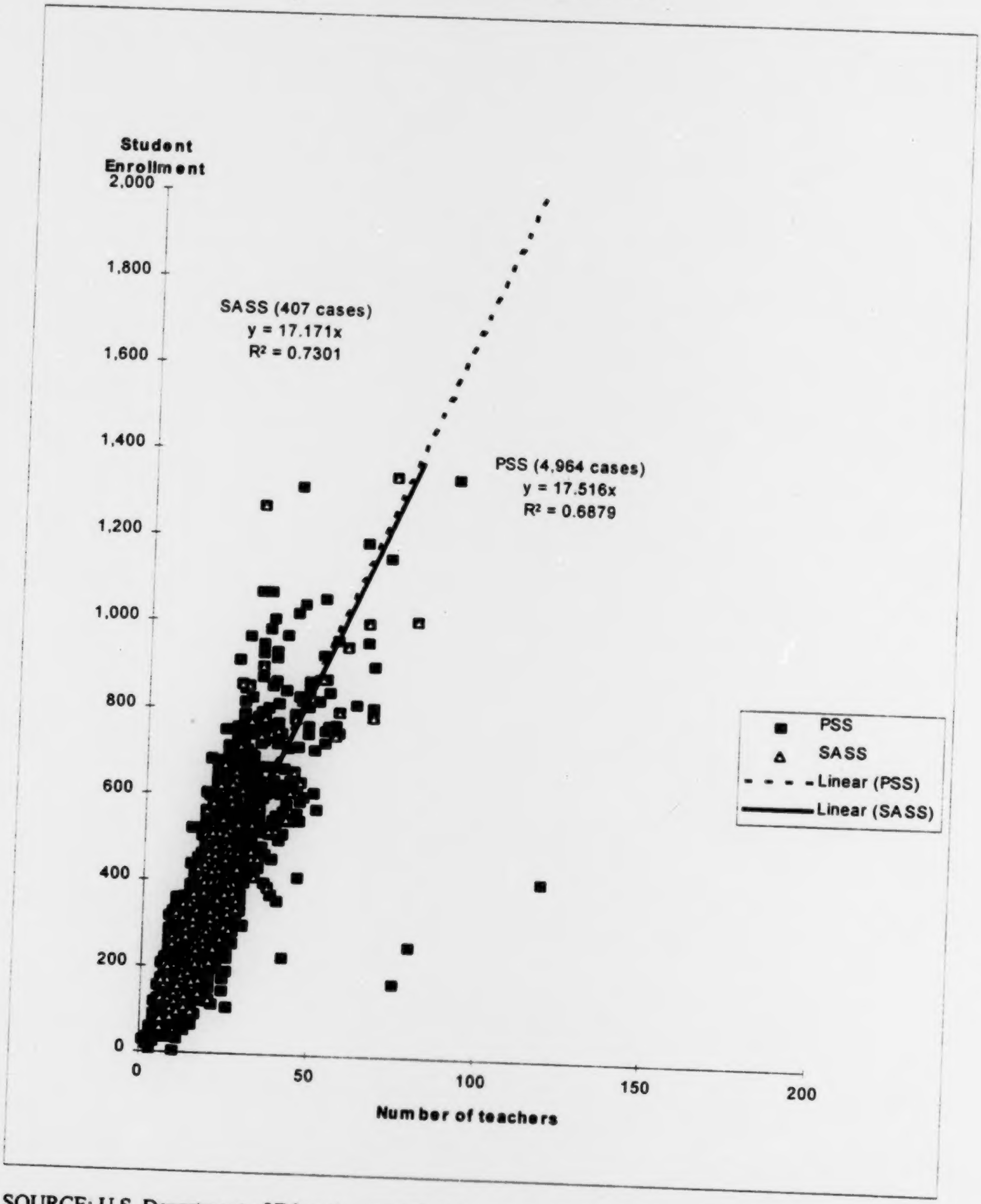
Table 1.2 -- Catholic Parochial: Weighted schools totals after excluding outliers (Based on 4,931 PSS and 401 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	5,093	5,061	32
Teachers	77,909	75,719	2,190
Students	1,378,215	1,352,296	25,919

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 1.1 -- Catholic Parochial: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

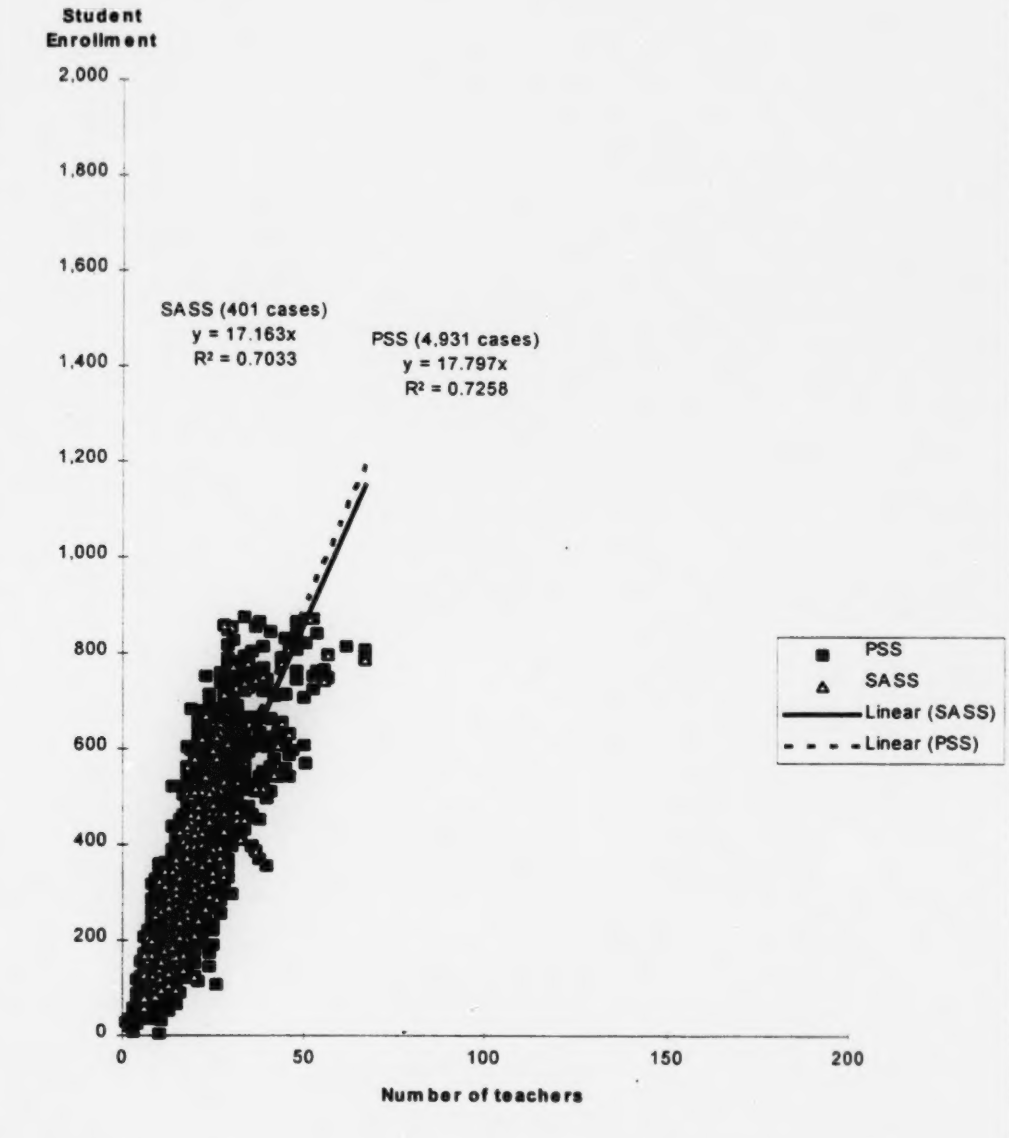
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 1.2 -- Catholic Parochial: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

values for the student/teacher ratios in both (unweighted) samples are nearly the same too -- at $R^2 = .69$ (PSS) and $R^2 = .73$ (SASS).

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing this new estimator, a decision was first made about which sample cases to use (see section 3.1.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.1.2). The results of the basic GLS were also obtained (section 3.1.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and Basic GLS versions (section 3.1.4). An independent assessment (section 3.1.5) concludes the discussion.

3.1.1 Determining Outliers. -- Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 1.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

For the Catholic Parochial typology, simple visual inspection seemed sufficient, resulting in a reduced PSS sample(from 4,964 to 4,931 cases)and a correspondingly reduced SASS sample (from 407 to 401 cases). Figure 1.2 is the plot of the remaining cases. Notice that the student/teacher relationships are little changed overall from those in figure 1.1; however, the scatter in both samples is considerably tighter.

3.1.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. The starting point was the new typology totals for PSS and SASS. These are shown in table 1.2 below.

To carry out the Olkin GLS, the schools were grouped into three school size classes (under 150 students, 150 to 499, and 500 and above). After the Olkin adjustment to each of the three school size groups, the difference between PSS and SASS had shrunk considerably in overall absolute value to

$$\underline{\mathbf{d}} = \begin{matrix} -8 \\ 707 \\ -9031 \end{matrix}$$

The matrix M was obtained by tabulating the 1993-94 SASS file for the Catholic Parochial schools in the SASS sample. The values are

$$\begin{matrix} 401 & 6674 & 116836 \\ 6674 & 140180 & 2405940 \\ 116836 & 2405940 & 44352886 \end{matrix}$$

Solving for $\underline{\lambda}$ yields

$$\underline{\lambda}' = (-0.3129, +0.13469, -0.006686)$$

and the Olkin GLS weights are of the form

$$u_i = w_i - 0.3129 + 0.13469t_i - 0.006686s_i$$

Notice that all the original weights are lowered (by about .3); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered further (usually they would not be "lowered further" except for the very largest schools). These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts are so small. One final comment, while the values for $\underline{\lambda}$ are only shown to six significant digits, the calculations have been carried out in double precision.

3.1.3 Basic GLS Procedure Employed. -- To carry out the basic GLS weighting the equation

$$\underline{\lambda} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 1.2 above that $\underline{\mathbf{d}}$ for the basic GLS would be

$$\underline{\mathbf{d}} = \begin{matrix} 31 \\ 2,190 \\ 25,919 \end{matrix}$$

The matrix **M** is again obtained by tabulating the eligible portion of the 1993-94 SASS file for the Catholic Parochial schools in the SASS sample. Because the Olkin and Basic GLS employ just the unweighted sample to calculate **M**, it is the same for both (and hence not shown).

Solving for $\underline{\lambda}$ yields this time

$$\underline{\lambda}' = (-0.78907, +0.10853, -0.00322)$$

and the basic GLS weights are of the form

$$u_i = w_i - 0.78907 + 0.10853t_i - 0.00322s_i$$

Notice that all the original weights are again lowered (but by over twice as much this time as was done for the Olkin GLS); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered further (usually they would not be "lowered further" except for the very largest schools). These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts remain, as was the case with the Olkin GLS, quite small. Looking just at the equation, concerns about negative weights might arise but, as will be seen below, these did not materialize.

3.1.4 Operational Characteristics. -- Both the Basic and Olkin GLS reweighting done, as described above, seems to have worked well. To indicate why this observation is made, several "diagnostics" will be looked at. One statistic that may merit immediate attention is what happened to the spread in the weights themselves. Did the spread grow larger or smaller?

Figure 1.3 provides this information in its upper panels, which compare the original and two GLS adjustments. Both GLS weights have a smaller spread than does the original SASS weight. This follows by noting that the original (or y) weights in the upper panel are slightly larger than the GLS (or x) weights; obviously, though, these differences are in no way important.

The R^2 values shown in the upper panel in figure 1.3 might be commented on too, along with the appearance of the scatter itself. In particular, very little evidence exists to indicate that either of the reweighting approaches altered the original weights in any major way. The R^2 values are both above 0.97 and most of the points lie close to the 45 degree line. The problem of negative weights did not arise either and there was only one case where the weight was less than one.

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 1.3 will continue to be our source. This time, though, look at panel in the second row of the scatterplot matrix where the Basic and Olkin GLS are graphed. The Basic GLS (or y) values are slightly larger than the Olkin GLS (or x) values; but ever so slightly. There is virtually no difference in the weights -- as evidenced by an R^2 of .99 between the two methods. The plotted points confirm this.

3.1.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Catholic Parochial Typology are available in tables 1.3 and 1.4, plus figure 1.4:

-- Table 1.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.

-- Table 1.4 is based on table 1.3 but focuses directly on percentage differences between the three SASS estimates and PSS.

-- Figure 1.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Notice that a log scale has been used here.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 1.3 and especially 1.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does the best on the average. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS in 11 out of 18 times and closer than the Basic GLS in 9 out of 18 comparisons. Thus in half or more of the cases, the Olkin method is to be favored. The data by community type are more mixed, as might be expected since the Olkin approach did not try to control by community type, as it had by school size.

In figure 1.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already very close to the PSS, since the best fit regression equation which connects the various estimates is $y = .9998x$. There is some roughness around this average, however, as displayed visually and summarized by the R^2 value which is equal to $R^2 = .9532$. Still this is quite good, suggesting that the SASS sample of Catholic Parochial schools is excellent.

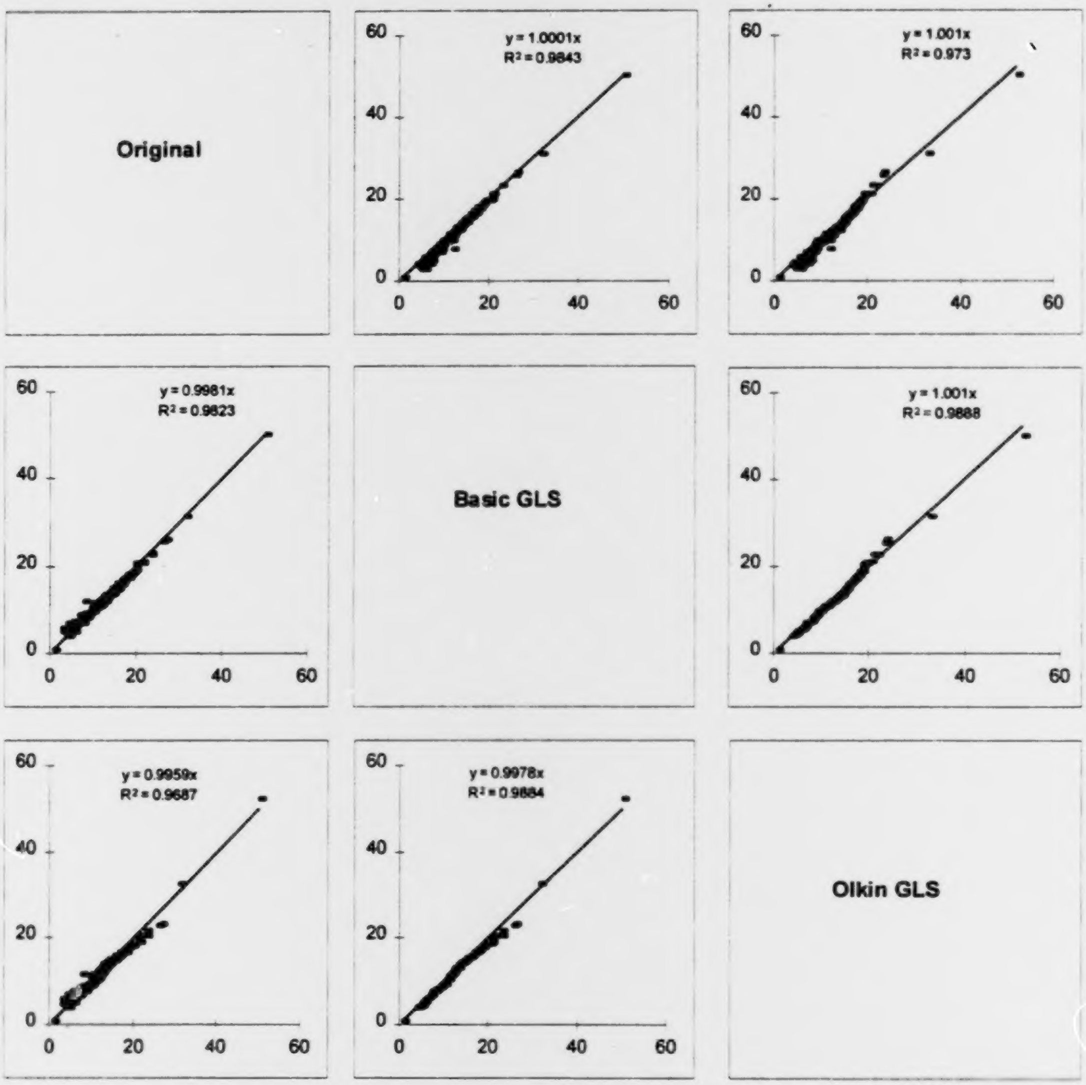
As in figure 1.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit yields the relationship $y = .9938x$. Again, the average results for this method remain good. Somewhat less roughness is exhibited around the average as evidenced by the slightly larger R^2 value in this case($R^2 = .9609$).

Finally, in figure 1.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields a relationship with the PSS totals of $y = .9977x$. The average results for this method are slightly better than the other two. In particular, somewhat less roughness is exhibited around the average as evidenced by the slightly larger R^2 value in this case($R^2 = .9634$).

What can be concluded about this typology? Neither the Basic nor the Olkin GLS methods seem in any way inferior overall to the original SASS weighted file. To their credit, they both hit the overall PSS school, teacher, and student totals exactly. The Olkin method, furthermore and not surprisingly, does as well or better than the other two when estimates are looked at by school size.

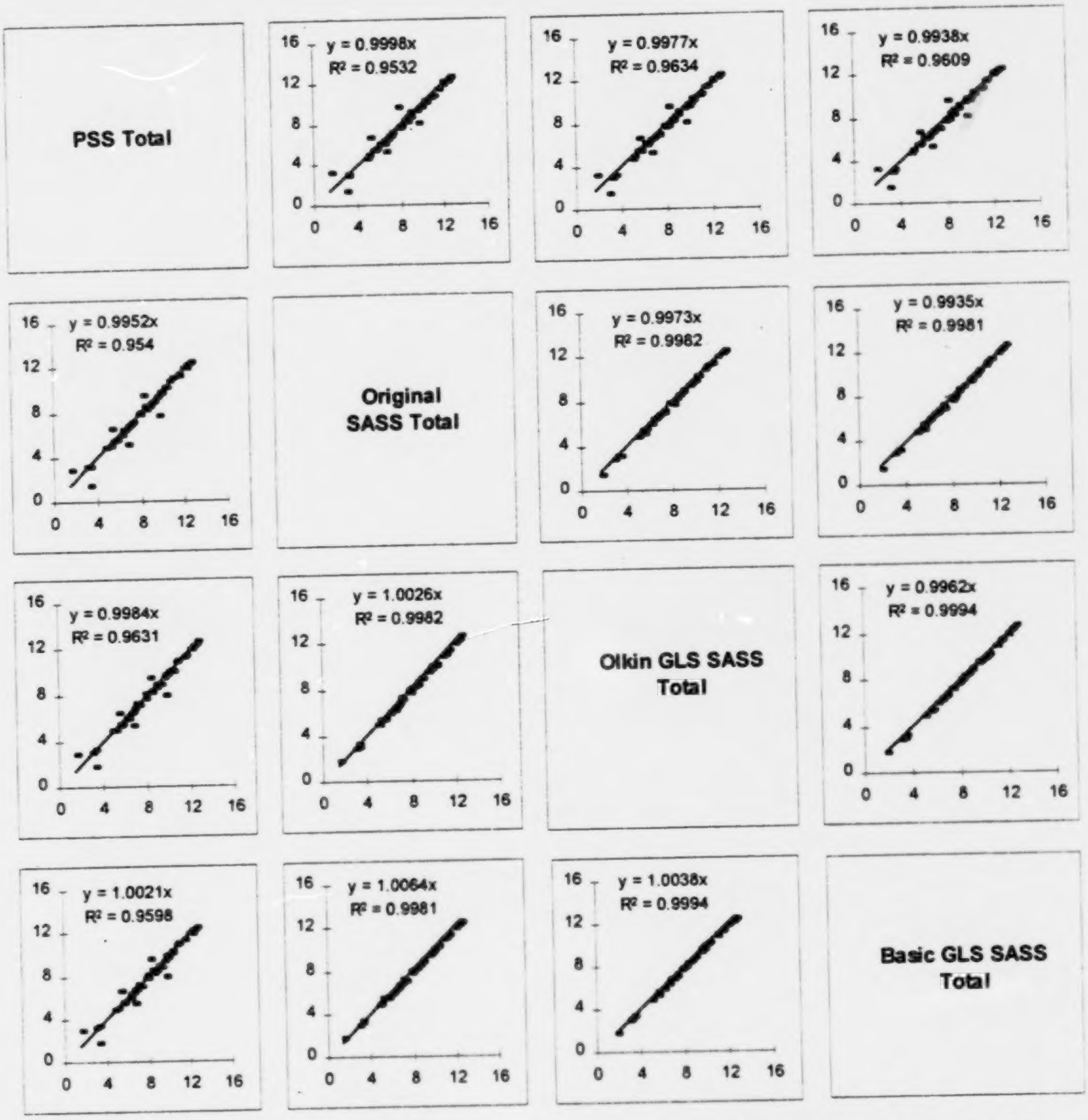
While the overall GLS performance is good to quite good, nonetheless, in individual cells, especially for the largest schools, the GLS estimates seem to have made matters worse. Ways to have done better were possible and, in typologies adjusted later, some were tried. In the summary and recommendations section, comments will be made about how the Olkin GLS might be improved further, leading to still better results.

Figure 1.3 – Catholic Parochial: Scatterplot matrix comparing original, basic GLS, and Olkin GLS weights



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993- 94.

Figure 1.4 – Catholic Parochial: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and basic GLS SASS totals by school size and community type from Table 1.3
(in log scale)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

BEST COPY AVAILABLE

Table 1.3 – Catholic Parochial: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part I - PSS total (4,931 schools)					
1 - 149	School	222	261	458	941
	Teacher	2,103	2,311	3,849	8,262
	Student	25,608	27,713	46,694	100,015
150 - 299	School	1,227	870	432	2,528
	Teacher	15,644	11,900	5,673	33,217
	Student	274,641	194,800	89,701	559,142
300 - 499	School	609	454	104	1,168
	Teacher	11,795	9,168	2,242	23,205
	Student	233,001	171,599	39,600	444,200
500 - 749	School	236	149	25	409
	Teacher	6,465	4,088	767	11,319
	Student	137,713	85,511	14,293	237,517
750 +	School	24	18	4	46
	Teacher	982	736	189	1,907
	Student	19,121	14,883	3,337	37,341
Total	School	2,318	1,752	1,022	5,092
	Teacher	36,988	28,202	12,719	77,909
	Student	690,084	494,507	193,624	1,378,215
Part II - Original SASS total (401 schools)					
1 - 149	School	216	261	563	1,041
	Teacher	2,421	2,277	4,813	9,511
	Student	28,262	25,000	60,574	113,836
150 - 299	School	1,238	865	393	2,496
	Teacher	15,196	11,868	5,126	32,190
	Student	274,690	198,691	84,185	557,565
300 - 499	School	596	372	118	1,086
	Teacher	11,530	6,992	2,738	21,260
	Student	228,439	138,177	45,032	411,648
500 - 749	School	238	135	4	377
	Teacher	6,617	3,595	153	10,366
	Student	142,174	75,851	2,034	220,059
750 +	School	22	22	17	61
	Teacher	1,003	754	634	2,392
	Student	18,046	17,861	13,280	49,187
Total	School	2,310	1,656	1,095	5,061
	Teacher	36,768	25,487	13,465	75,719
	Student	691,611	455,580	205,106	1,352,296

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 1.3 – Catholic Parochial: Estimates by school size and community type, PSS and SASS compared (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part III - Olkin GLS SASS total (401 schools)					
1 - 149	School	197	236	502	935
	Teacher	2,230	2,079	4,316	8,625
	Student	25,707	22,669	54,017	102,393
150 - 299	School	1,280	902	411	2,593
	Teacher	15,818	12,413	5,394	33,626
	Student	283,942	206,702	87,796	578,440
300 - 499	School	610	383	129	1,121
	Teacher	11,906	7,254	3,035	22,195
	Student	233,284	142,014	49,232	424,530
500 - 749	School	239	135	6	380
	Teacher	6,872	3,683	216	10,772
	Student	143,128	75,890	2,868	221,886
750 +	School	26	20	17	63
	Teacher	1,321	737	634	2,691
	Student	21,044	16,874	13,048	50,966
Total	School	2,352	1,676	1,064	5,092
	Teacher	38,148	26,167	13,594	77,909
	Student	707,105	464,149	206,961	1,378,215
Part IV - Basic GLS SASS total (401 schools)					
1 - 149	School	217	260	557	1,034
	Teacher	2,447	2,280	4,781	9,507
	Student	28,328	24,911	59,962	113,201
150 - 299	School	1,226	863	392	2,482
	Teacher	15,150	11,886	5,147	32,183
	Student	272,311	198,251	83,966	554,527
300 - 499	School	603	378	127	1,107
	Teacher	11,773	7,162	3,009	21,944
	Student	231,149	140,305	48,892	420,346
500 - 749	School	251	141	6	398
	Teacher	7,240	3,826	223	11,288
	Student	150,916	79,020	2,953	232,889
750 +	School	28	24	19	71
	Teacher	1,415	870	702	2,987
	Student	22,965	19,784	14,502	57,251
Total	School	2,325	1,665	1,102	5,092
	Teacher	38,025	26,023	13,862	77,909
	Student	705,668	462,271	210,275	1,378,215

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

BEST COPY AVAILABLE

Table 1.4 – Catholic Parochial: Estimates by school size and community type, PSS and SASS compared in percent difference

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent Difference from PSS and Original SASS totals					
1 - 149	School	2.80	-0.19	-23.05	-10.60
	Teacher	-15.13	1.45	-25.05	-15.12
	Student	-10.36	9.79	-29.73	-13.82
150 - 299	School	-0.94	0.60	8.83	1.26
	Teacher	2.86	0.27	9.64	3.09
	Student	-0.02	-2.00	6.15	0.28
300 - 499	School	2.24	17.98	-13.20	6.99
	Teacher	2.24	23.74	-22.13	8.38
	Student	1.96	19.48	-13.72	7.33
500 - 749	School	-0.78	8.82	84.01	7.81
	Teacher	-2.35	12.05	79.98	8.42
	Student	-3.24	11.30	85.77	7.35
750 +	School	8.20	-17.74	-313.98	-30.90
	Teacher	-2.14	-2.52	-236.23	-25.45
	Student	5.62	-20.01	-298.00	-31.72
Total % diff from PSS	School	0.37	5.49	-7.19	0.61
	Teacher	0.60	9.63	-5.86	2.81
	Student	-0.22	7.87	-5.93	1.88
Percent Difference from PSS and Olkin GLS SASS totals					
1 - 149	School	11.41	9.56	-9.74	0.61
	Teacher	-6.05	10.02	-12.14	-4.39
	Student	-0.39	18.20	-15.68	-2.38
150 - 299	School	-4.35	-3.63	4.77	-2.54
	Teacher	-1.12	-4.32	4.93	-1.23
	Student	-3.39	-6.11	2.12	-3.45
300 - 499	School	-0.02	15.67	-23.71	3.97
	Teacher	-0.94	20.87	-35.38	4.35
	Student	-0.12	17.24	-24.32	4.43
500 - 749	School	-1.34	8.82	77.48	7.10
	Teacher	-6.30	9.89	71.78	4.84
	Student	-3.93	11.25	79.94	6.58
750 +	School	-8.37	-11.10	-306.51	-36.13
	Teacher	-34.52	-0.09	-235.81	-41.15
	Student	-10.06	-13.38	-291.03	-36.49
Total % diff from PSS	School	-1.44	4.32	-4.14	0.00
	Teacher	-3.13	7.21	-6.88	0.00
	Student	-2.47	6.14	-6.89	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 1.4 – Catholic Parochial: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	% diff from PSS
		Percent difference from PSS and Basic GLS SASS totals			
1 - 149	School	2.46	0.47	-21.78	-9.88
	Teacher	-16.34	1.34	-24.22	-15.07
	Student	-10.62	10.11	-28.42	-13.18
150 - 299	School	0.02	0.76	9.07	1.82
	Teacher	3.16	0.11	9.28	3.11
	Student	0.85	-1.77	6.39	0.83
300 - 499	School	1.13	16.86	-22.48	5.15
	Teacher	0.19	21.88	-34.22	5.43
	Student	0.79	18.24	-23.47	5.37
500 - 749	School	-6.51	5.25	76.83	2.78
	Teacher	-11.99	6.41	70.94	0.27
	Student	-9.59	7.59	79.34	1.95
750 +	School	-18.08	-30.03	-351.81	-52.66
	Teacher	-44.09	-18.19	-272.29	-56.68
	Student	-20.10	-32.93	-334.62	-53.32
Total % diff from PSS	School	-0.30	4.95	-7.79	0.00
	Teacher	-2.80	7.73	-8.98	0.00
	Student	-2.26	6.52	-8.60	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

3.2 CATHOLIC DIOCESAN TYPOLOGY

The Catholic Diocesan typology represents the second largest type of private school in terms of students. For example, in the 1993-94 Private School Survey, there were an estimated 2,371 Catholic Diocesan schools or about 9% of the private school total for that year. However, the number of students in such schools, at over 750,000 is a much larger percentage of the overall student total (at 16 percent).

In table 2.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the SASS school totals are greater than the PSS (by about 2.6%); SASS also estimates more teachers and students than are shown in PSS (3.1% and 5.2% more respectively).

To set the stage for the calculations that follow, it might be worth looking at figure 2.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black □'s. Overlaying the PSS data and shown as gray Δ's are the data from the 262 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS and along essentially the same axis. In fact, the slope of the student/teacher relationship is 16.7 for PSS and 16.4 for SASS -- virtually indistinguishable (Indeed, the least squares lines are touching over most of their length). While not directly comparable because of differences in sample designs, the R^2 values for the student/teacher ratios in both (unweighted) samples show analytically what can be observed from the graph, namely that the scatter is quite close around the average teacher/student relationship. These values are $R^2 = .84$ (PSS) and $R^2 = .88$ (SASS).

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing these new GLS estimators, a decision was first made about which sample cases to use (see section 3.2.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.2.2). The results of the Basic GLS were also obtained (section 3.2.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and Basic GLS versions (section 3.2.4). An independent assessment (section 3.2.5) concludes the discussion.

3.2.1 Determining Outliers. -- Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 2.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

Table 2.1 -- Catholic Diocesan: Weighted schools totals before excluding outliers (Based on 2,285 PSS and 262 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	2,371	2,432	-61
Teachers	44,997	46,400	-1,402
Students	751,175	790,442	-39,267

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

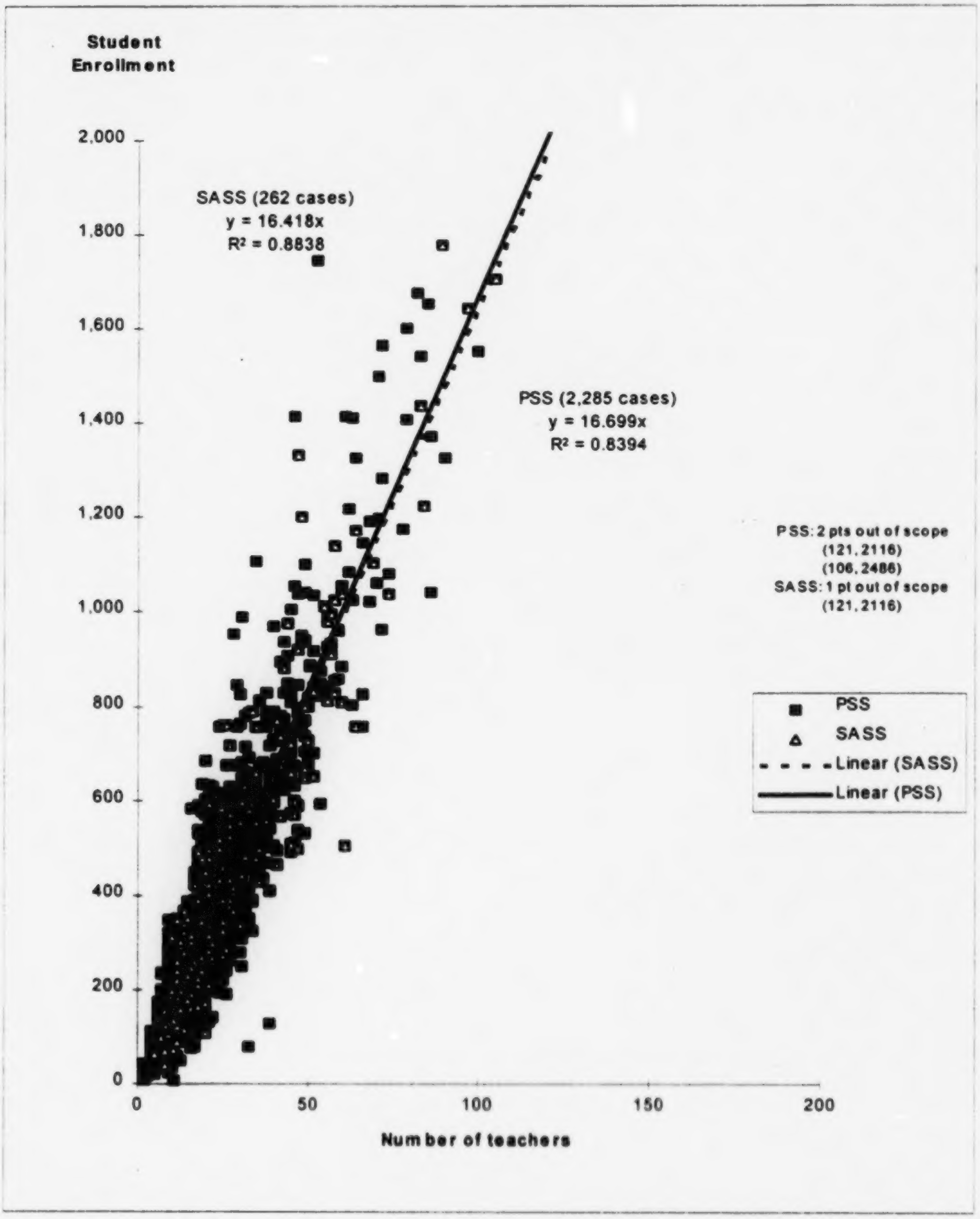
Table 2.2 -- Catholic Diocesan: Weighted schools totals after excluding outliers (Based on 2,262 PSS and 256 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	2,347	2,424	-77
Teachers	43,112	44,948	-1,836
Students	713,845	762,146	-48,301

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 2.1 – Catholic Diocesan: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

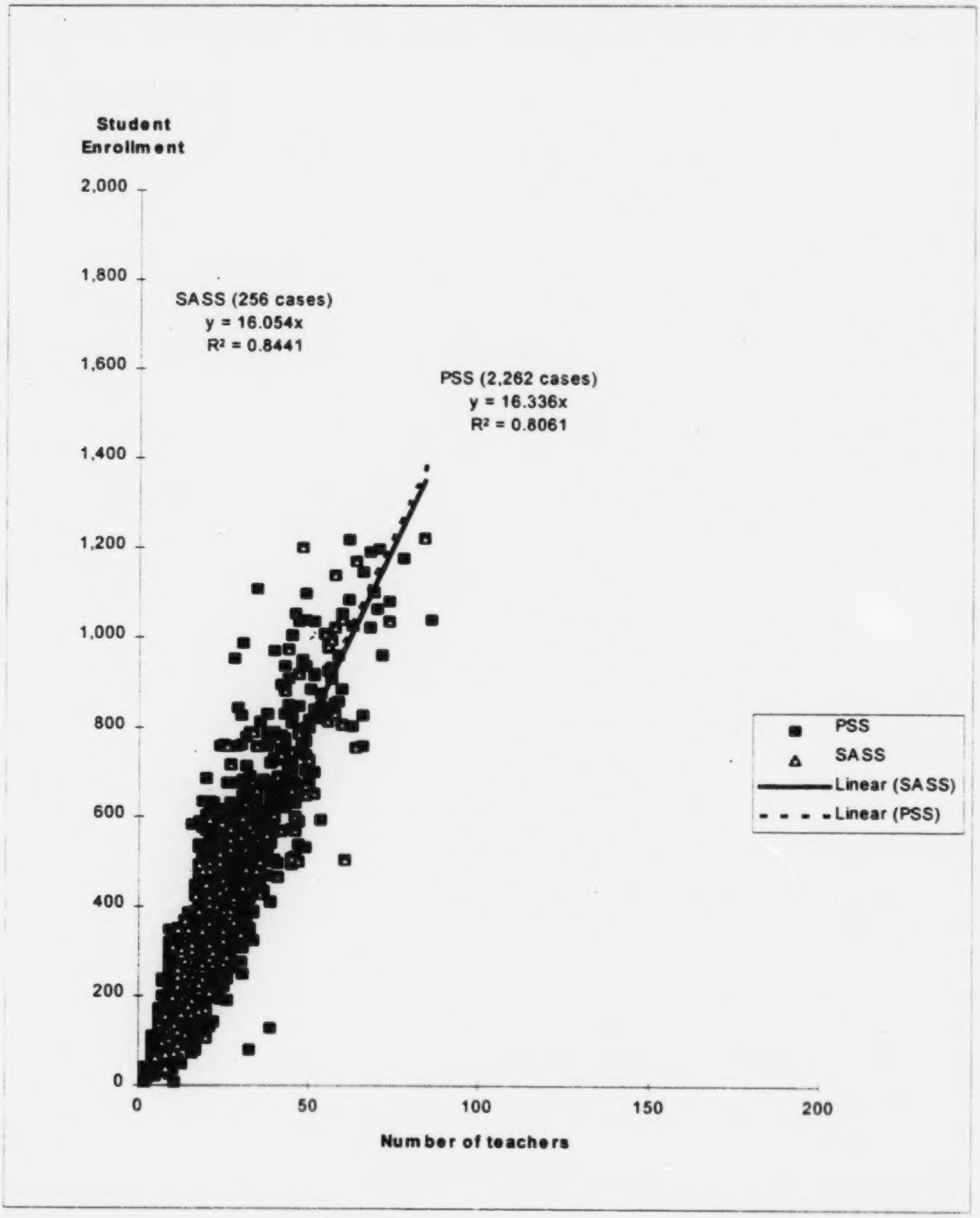
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 2.2 – Catholic Diocesan: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

BEST COPY AVAILABLE

For the Catholic Diocesan typology, simple visual inspection seemed sufficient, resulting in a reduced PSS sample (from 2,285 to 2262 cases) and a correspondingly reduced SASS sample (from 262 to 256 cases). Figure 2.2 is the plot of the remaining cases. Notice that the student/teacher relationships are little changed overall from those in figure 2.1; however, the scatter of points in both samples is visually even tighter. (Ironically, because the largest (influential) observations have been dropped, the R^2 values dropped slightly.)

3.2.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. The starting point was the new typology totals for PSS and SASS. These are shown in table 2.2 below.

To carry out the Olkin GLS, the schools were placed into four school size classes (under 150, 150 to 499, 500 to 749, 750 and above).

After the Olkin adjustment, the difference between the PSS and SASS estimates had shrunk considerably (in overall absolute value) to

$$\underline{\mathbf{d}} = \begin{matrix} 12 \\ 289 \\ -7974 \end{matrix}$$

The matrix M was obtained by tabulating the 1993-94 SASS file for the Catholic Diocesan schools in the SASS sample. The values are

$$\begin{matrix} 256 & 6338 & 103863 \\ 6338 & 218042 & 3500414 \\ 103863 & 3500414 & 58790303 \end{matrix}$$

Solving for $\underline{\boldsymbol{\lambda}}$ yields

$$\underline{\boldsymbol{\lambda}}' = (+.19973, +0.07611, -0.00502)$$

and the Olkin GLS weights are of the form

$$u_i = w_i + 0.19973 + 0.07611t_i - 0.00502s_i$$

Notice that all the original weights are raised slightly (by about .2); and, then, depending on the teacher and student counts in the sampled school, they may be in-creased again or lowered further (usually they would not be "lowered further" except for the very largest schools). These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts are so small. One final comment, while the values for $\underline{\lambda}$ are only shown to six significant digits, the calculations have been carried out in double precision.

3.2.3 Basic GLS Procedure Employed. -- To carry out the basic GLS weighting the equation

$$\underline{\lambda} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 1.2 above that $\underline{\mathbf{d}}$ for the Basic GLS would be

$$\underline{\mathbf{d}} = \begin{matrix} -67 \\ -1836 \\ -48301 \end{matrix}$$

The matrix **M** is again obtained by tabulating the eligible portion of the 1993-94 SASS file for the Catholic Diocesan schools in the SASS sample. Because the Olkin and Basic GLS employ just the unweighted sample to calculate **M**, it is the same for both (and hence not shown).

Solving for $\underline{\lambda}$ yields this time

$$\underline{\lambda}' = (+.02346, +0.10765, -0.00727)$$

and the basic GLS weights are of the form

$$u_i = w_i + 0.02346 + 0.10765t_i - 0.00727s_i$$

Notice that all the original weights are raised very slightly (not as much as the Olkin GLS though); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered further (usually they would not be "lowered further" except for the very largest schools). These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts are so small. They are larger in absolute value, though, than for the Olkin adjustment, a pattern that was expected (and which turns out to be generally true for all typologies). Looking just at the equation, concerns about negative weights arise but, as will be seen below, these did not materialize.

3.2.4 Operational Characteristics. -- To examine the Basic and Olkin GLS reweighting done, several "diagnostics" will be looked at. One statistic that may merit immediate attention is what happened to the spread in the weights themselves. Did the spread grow larger or smaller?

Figure 2.3 provides this information in its upper panels, which compare the original and two GLS adjustments. The Basic GLS weight has a smaller spread than does the original SASS weight. This follows by noting that the original (or y) weights in the upper panel are slightly larger than the GLS (or x) weights (since the equation which fits them is $y = 1.017x$). For the Olkin GLS, the variability in the weights is somewhat greater than in the original SASS (with the equation relating them being of the form $y = .994x$).

While the overall differences in scale between the weights appear unimportant, the scatter for the Olkin GLS shows a distinct break between the original data and the final Olkin weights for the largest schools.

The R^2 values shown in the upper panel in figures 2.3 might be commented on too. Despite the appearance of the scatter itself, very little evidence exists to indicate that either of the reweighting approaches altered the original weights in any major way. The R^2 values are both at or above 0.92 and most of the points lie close to the 45 degree line. The problem of negative weights did not arise for the Olkin GLS, although there was one school with a weight smaller than one. For the Basic GLS, the results were not quite as good. A negative weight existed, and there were two cases with weights less than one.

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 2.3 will continue to be our source. This time, though, look at panel in the second row of the scatterplot matrix where the Basic and Olkin GLS are graphed. The basic GLS (or y) values are slightly smaller than the Olkin GLS (or x) values; but ever so slightly. There is virtually no overall difference in the weights -- as evidenced by an R^2 of .95 between the two methods. The plotted points do indicate some departures though, as noted earlier, among the largest schools.

3.2.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Catholic Diocesan Typology are available in tables 2.3 and 2.4, plus figure 2.4:

-- Table 2.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.

-- Table 2.4 is based on table 2.3 but focuses directly on percentage differences between the three SASS estimates and PSS.

-- Figure 2.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Notice that a log scale has been used here.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 2.3 and especially 2.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does the best on the average. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS 11/18 times and closer than the Basic GLS in 13/18 comparisons. Thus, in well over half of the cases, the Olkin method is to be favored.

The results by community type are more mixed, as might be expected, since the Olkin approach did not try to control by community type (as it had by school size). The rural schools estimates, for example, were better using the original SASS weights than with either of the GLS estimators. On the other hand, the Olkin GLS was marked better than the original SASS data for central city or urban fringe/small town estimates.

In figure 2.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already very close to the PSS, since the best fit regression equation which connects the various estimates is $y = 1.0005x$. There is a little roughness around this average, however, as displayed visually and summarized by the R^2 value which is equal to $R^2 = .9848$. This is extremely good, suggesting that the SASS sample of Catholic Diocesan schools is excellent.

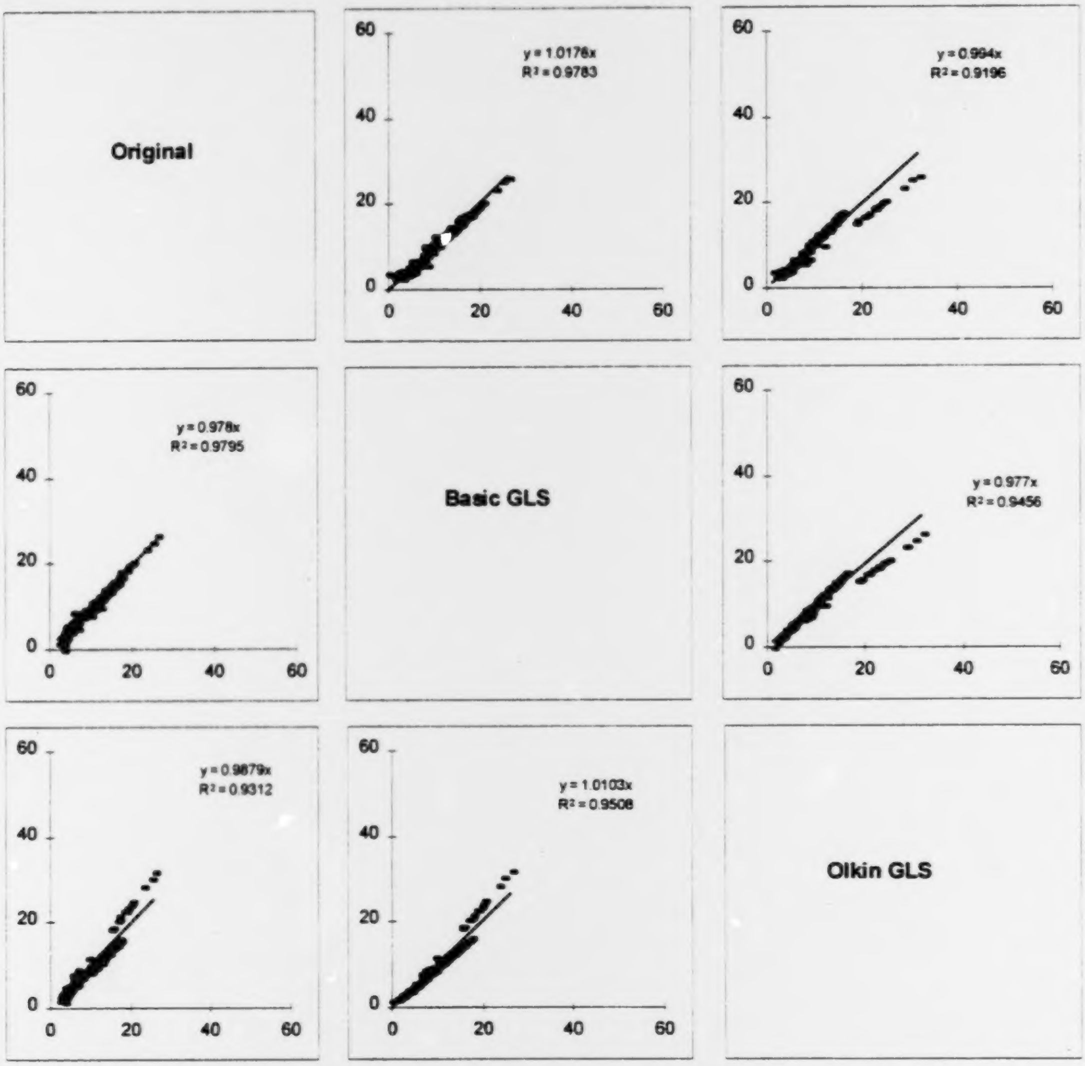
As in figure 2.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit is similar, yielding the relationship $y = 1.0026x$. Again, the average results for this method remain good. Somewhat less roughness is exhibited around the average as evidenced by the slightly larger R^2 value in this case($R^2 = .987$).

Finally, in figure 2.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields the relationship $y = 1.0073x$. The average results for this method are again comparable to the other two, with an R^2 value in this case of $R^2 = .9842$.

What can be concluded about this typology? Neither the Basic nor the Olkin GLS methods seem in any way inferior overall to the original SASS weighted file. To their credit, they both hit the overall PSS school, teacher, and student totals exactly. The Olkin method, furthermore and not surprisingly, does as well or better than the other two when estimates are looked at by school size. The Basic GLS method has a negative weight; and, if it were to be used, further work would be needed on it.

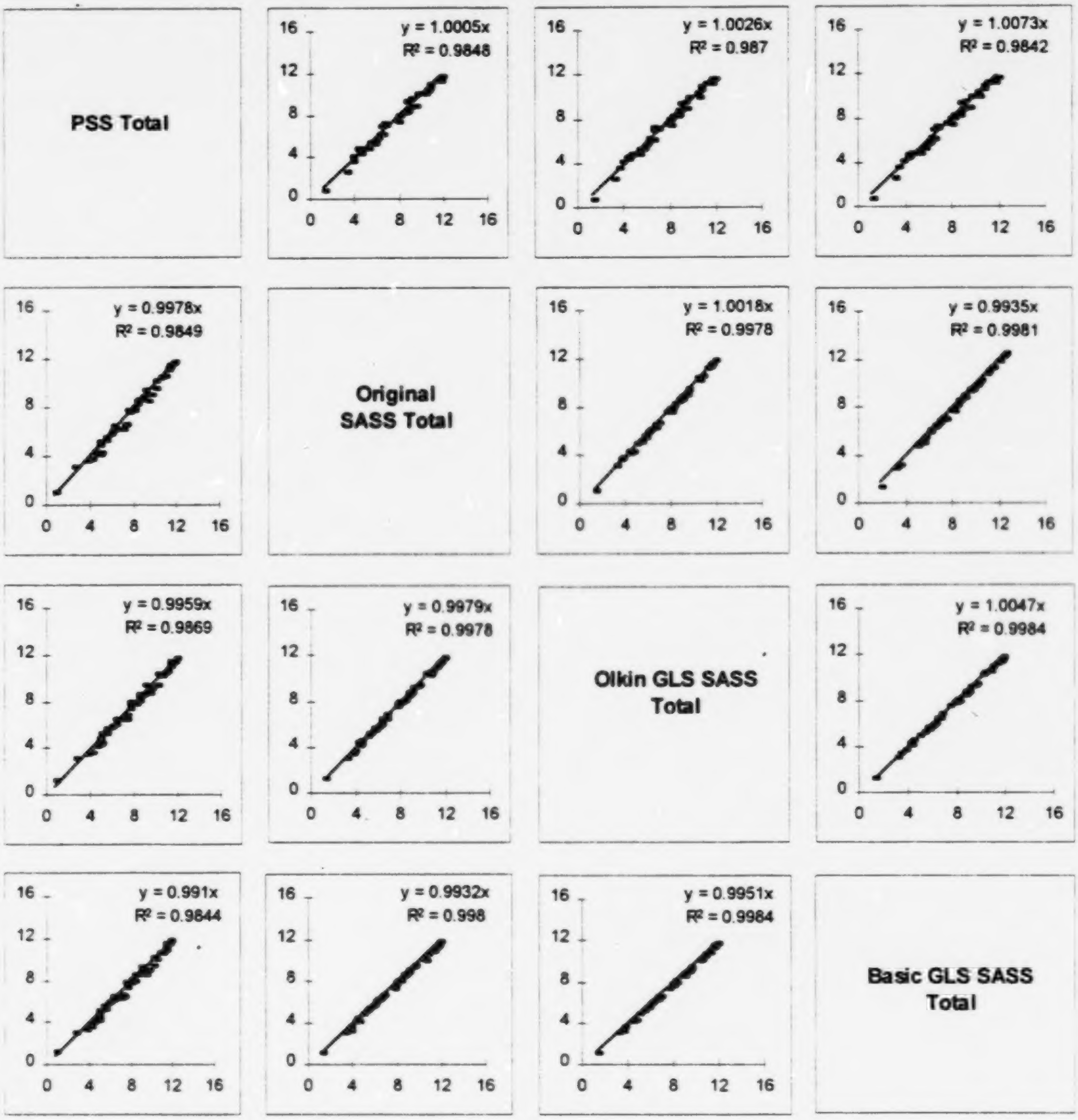
In the summary and recommendations section, additional comments will be made about how the Olkin GLS might be improved further, leading to still better results.

Figure 2.3 – Catholic Diocesan: Scatterplot matrix comparing original, basic GLS, and Olkin GLS weights



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 2.4 – Catholic Diocesan: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and basic GLS SASS totals by school size and community type from Table 2.3
(in log scale)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 2.3 -- Catholic Diocesan: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part I - PSS total (2,262 schools)					
1 - 149	School	126	112	205	443
	Teacher	1,202	968	1,751	3,921
	Student	13,305	10,689	21,372	45,366
150 - 299	School	474	344	202	1,021
	Teacher	6,417	4,815	2,922	14,154
	Student	105,817	77,506	42,187	225,510
300 - 499	School	302	193	57	552
	Teacher	6,691	4,067	1,333	12,092
	Student	115,864	73,173	21,262	210,299
500 - 749	School	117	88	12	217
	Teacher	3,904	2,888	443	7,235
	Student	69,818	52,497	7,215	129,530
750 +	School	77	35	2	113
	Teacher	3,815	1,771	124	5,710
	Student	69,206	32,337	1,598	103,141
Total	School	1,096	772	479	2,347
	Teacher	22,029	14,509	6,574	43,112
	Student	374,010	246,201	93,634	713,845
Part II - Original SASS total (256 schools)					
1 - 149	School	73	61	229	363
	Teacher	724	498	1,988	3,210
	Student	7,985	5,183	25,609	38,777
150 - 299	School	461	411	186	1,058
	Teacher	6,245	5,399	2,953	14,597
	Student	105,486	89,506	40,061	235,053
300 - 499	School	345	271	41	658
	Teacher	7,047	5,626	701	13,375
	Student	133,502	103,563	13,611	250,676
500 - 749	School	131	66	22	219
	Teacher	4,888	2,329	627	7,844
	Student	79,003	39,644	12,091	130,738
750 +	School	78	37	3	118
	Teacher	3,747	2,009	166	5,922
	Student	69,654	34,837	2,410	106,901
Total	School	1,087	845	482	2,414
	Teacher	22,652	15,861	6,435	44,948
	Student	395,631	272,732	93,783	762,146

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 2.3 -- Catholic Diocesan: Estimates by school size and community type, PSS and SASS compared (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part III - Olkin GLS SASS total (256 schools)					
1 - 149	School	90	75	281	445
	Teacher	896	615	2,456	3,967
	Student	9,819	6,380	31,398	47,597
150 - 299	School	427	381	178	985
	Teacher	5,808	5,033	2,842	13,684
	Student	97,645	82,956	38,270	218,871
300 - 499	School	316	247	38	601
	Teacher	6,574	5,161	645	12,380
	Student	122,385	94,225	12,289	228,899
500 - 749	School	129	65	21	214
	Teacher	4,925	2,361	611	7,896
	Student	77,369	39,086	11,537	127,992
750 +	School	67	31	3	101
	Teacher	3,274	1,729	181	5,185
	Student	59,752	28,103	2,632	90,487
Total	School	1,028	799	520	2,347
	Teacher	21,478	14,900	6,735	43,112
	Student	366,969	250,750	96,127	713,845
Part IV - Basic GLS SASS total (256 schools)					
1 - 149	School	74	62	233	369
	Teacher	744	510	2,038	3,293
	Student	8,124	5,291	25,957	39,372
150 - 299	School	456	407	191	1,053
	Teacher	6,207	5,375	3,050	14,633
	Student	104,258	88,487	41,017	233,762
300 - 499	School	331	259	39	629
	Teacher	6,895	5,403	667	12,966
	Student	128,188	98,571	12,677	239,436
500 - 749	School	126	63	20	210
	Teacher	4,836	2,336	596	7,767
	Student	75,126	38,407	11,245	124,778
750 +	School	58	25	3	87
	Teacher	2,838	1,442	174	4,454
	Student	51,668	22,306	2,523	76,497
Total	School	1,045	817	485	2,347
	Teacher	21,519	15,067	6,526	43,112
	Student	367,363	253,062	93,420	713,845

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 2.4 – Catholic Diocesan: Estimates by school size and community type, PSS and SASS compared in percent difference

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	

Percent difference from PSS and original SASS totals

School Size		Central City	Urban Fringe / Large Town	Rural / Small Town	Total % diff from PSS
1 - 149	School	42.23	45.95	-11.68	18.18
	Teacher	39.77	48.59	-13.56	18.13
	Student	39.98	51.51	-19.83	14.52
150 - 299	School	2.92	-19.17	7.79	-3.57
	Teacher	2.68	-12.13	-1.06	-3.13
	Student	0.31	-15.48	5.04	-4.23
300 - 499	School	-14.35	-40.28	27.44	-19.11
	Teacher	-5.32	-38.35	47.44	-10.61
	Student	-15.22	-41.53	35.98	-19.20
500 - 749	School	-12.07	25.39	-76.38	-0.57
	Teacher	-25.22	19.37	-41.62	-8.42
	Student	-13.16	24.48	-67.58	-0.93
750 +	School	-1.21	-7.53	-45.10	-3.94
	Teacher	1.79	-13.45	-33.69	-3.71
	Student	-0.65	-7.73	-50.83	-3.65
Total % diff from PSS	School	0.78	-9.42	-0.60	-2.86
	Teacher	-2.83	-9.32	2.10	-4.26
	Student	-5.78	-10.78	-0.16	-6.77

Percent difference from PSS and Olkin GLS SASS totals

School Size		Central City	Urban Fringe / Large Town	Rural / Small Town	Total % diff from PSS
1 - 149	School	28.71	33.34	-36.94	-0.54
	Teacher	25.47	36.48	-40.28	-1.17
	Student	26.20	40.32	-46.92	-4.92
150 - 299	School	10.09	-10.65	12.10	3.49
	Teacher	9.49	-4.53	2.74	3.33
	Student	7.72	-7.03	9.28	2.94
300 - 499	School	-4.73	-27.93	34.25	-8.82
	Teacher	1.75	-26.91	51.64	-2.39
	Student	-5.63	-28.77	42.20	-8.84
500 - 749	School	-10.21	26.51	-67.83	1.37
	Teacher	-26.17	18.26	-37.85	-9.15
	Student	-10.81	25.55	-59.90	1.19
750 +	School	13.16	10.27	-58.82	10.98
	Teacher	14.18	2.35	-46.01	9.20
	Student	13.66	13.09	-64.71	12.27
Total % diff from PSS	School	6.19	-3.43	-8.63	0.00
	Teacher	2.50	-2.69	-2.45	0.00
	Student	1.88	-1.85	-2.66	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 2.4 – Catholic Diocesan: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and Basic GLS SASS totals					
1 - 149	School	40.89	44.69	-13.28	16.75
	Teacher	38.11	47.26	-16.41	16.02
	Student	38.94	50.50	-21.46	13.21
150 - 299	School	3.97	-18.13	5.75	-3.13
	Teacher	3.28	-11.64	-4.38	-3.38
	Student	1.47	-14.17	2.77	-3.66
300 - 499	School	-9.77	-34.01	31.92	-13.95
	Teacher	-3.05	-32.85	49.96	-7.23
	Student	-10.64	-34.71	40.38	-13.86
500 - 749	School	-7.38	27.76	-63.52	3.64
	Teacher	-23.88	19.13	-34.62	-7.36
	Student	-7.60	26.84	-55.85	3.67
750 +	School	24.34	26.72	-51.96	23.68
	Teacher	25.62	18.56	-39.97	22.00
	Student	25.34	31.02	-57.91	25.83
Total % diff from PSS	School	4.63	-5.76	-1.30	0.00
	Teacher	2.31	-3.84	0.73	0.00
	Student	1.78	-2.79	0.23	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

BEST COPY AVAILABLE

3.3 CATHOLIC PRIVATE TYPOLOGY

The Catholic Private typology is a fairly small proportion of all private schools. For example, in the 1993-94 Private School Survey, there were an estimated 833 Catholic Private schools or over 3% of the private school total for that year.

In table 3.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the SASS school total is smaller than the PSS (by about 2.8%); SASS, however, estimates more teachers and students than are shown in PSS (2.8% and 3.6% more respectively).

To set the stage for the calculations that follow, it might be worth looking at figure 3.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black □'s. Overlaying the PSS data and shown as gray Δ's are the data from the 148 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS and along essentially the same axis. In fact, the slopes of the student/teacher relationship are close, albeit distinguishable, being 13.6 for PSS and 14.1 for SASS. While not directly comparable because of differences in sample designs, the R^2 values for the student/teacher ratios in both (unweighted) samples show analytically what can be observed from the graph, namely that the scatter is quite close around the average teacher/student relationship. These values are $R^2 = .77$ (PSS) and $R^2 = .74$ (SASS).

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing these new GLS estimators, a decision was first made about which sample cases to use (see section 3.3.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.3.2). The results of the basic GLS were also obtained (section 3.3.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and Basic GLS versions (section 3.3.4). An independent assessment (section 3.3.5) concludes the discussion.

3.3.1 Determining Outliers. -- Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 3.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

For the Catholic Private typology, simple visual inspection resulted in reducing the PSS sample by 15 cases -- with a corresponding reduction in the SASS sample of 4 cases. Figure

Table 3.1 – Catholic Private: Weighted schools totals before excluding outliers (Based on 788 PSS and 148 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	833	810	23
Teachers	25,145	25,852	-707
Students	327,097	338,641	-11,544

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

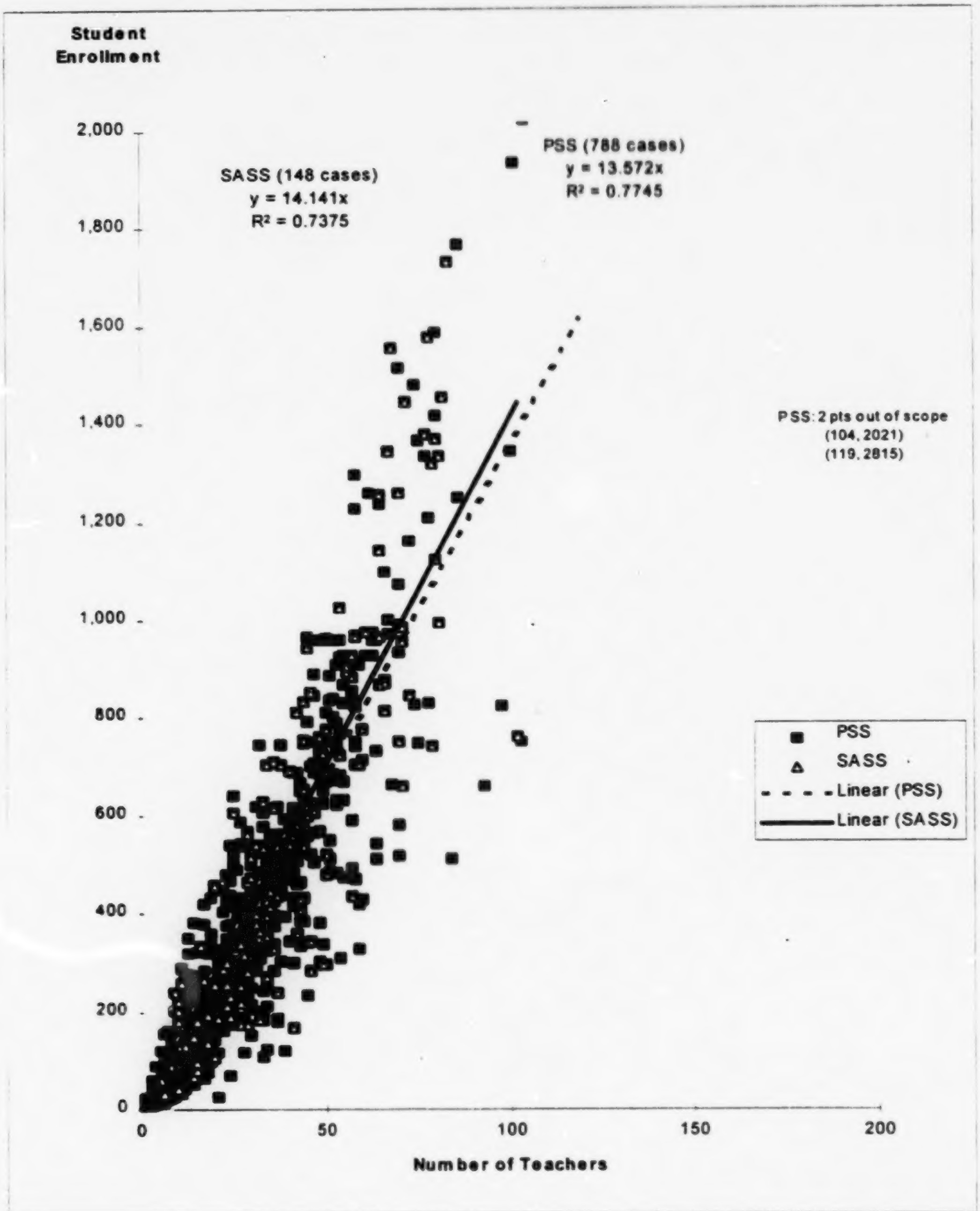
Table 3.2 – Catholic Private: Weighted schools totals after excluding outliers (Based on 733 PSS and 144 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	817	798	19
Teachers	23,724	24,894	-1,169
Students	304,702	322,275	-17,573

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 3.1 – Catholic Private: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

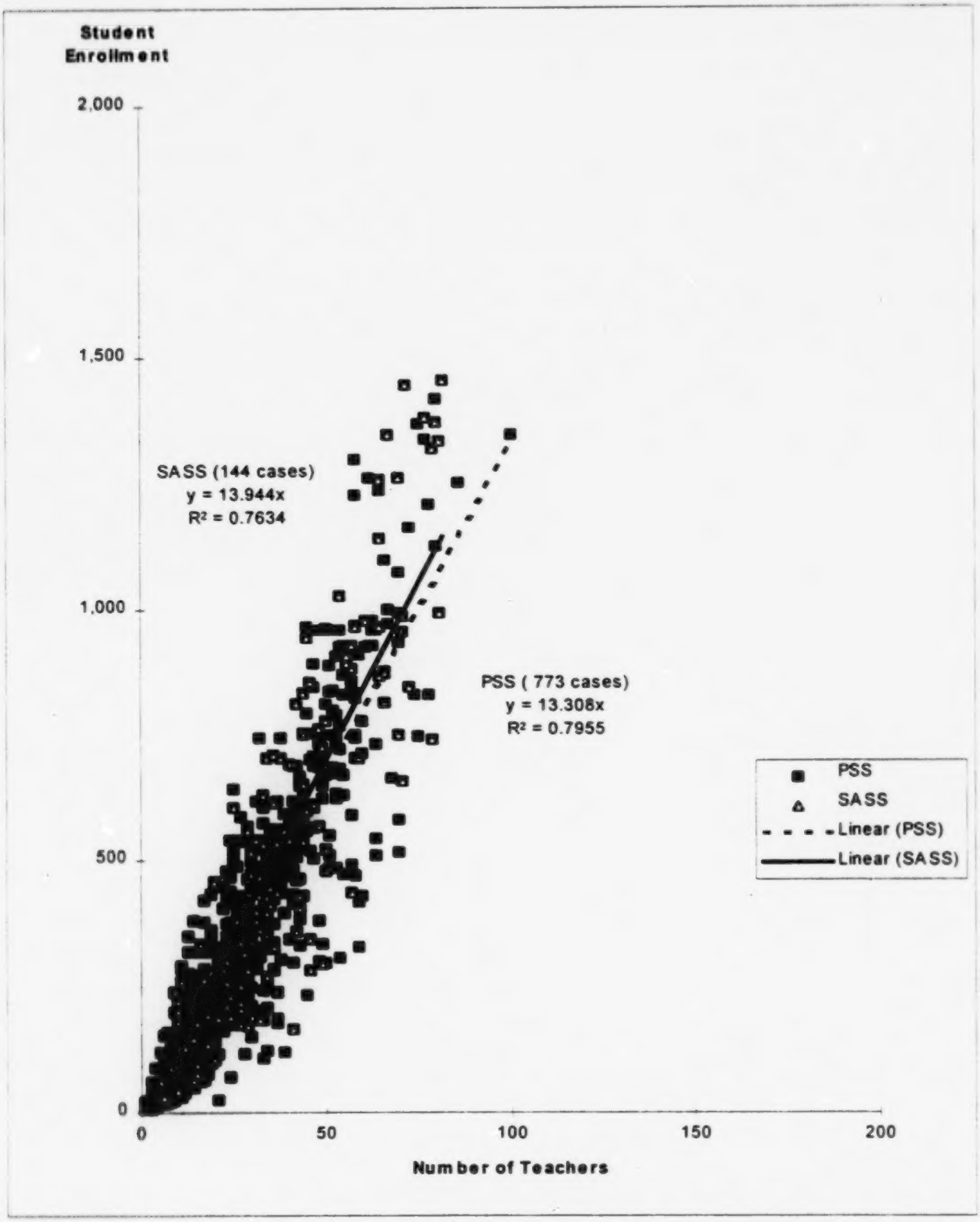
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 3.2 – Catholic Private: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys,1993-94.

BEST COPY AVAILABLE

3.2 is the plot of the remaining 773 PSS and 144 SASS cases. Notice that the student/teacher relationships are little changed overall from those in figure 3.1; however, the scatter of points in both samples is visually even tighter. The R^2 values reflect this, rising from .77 to .80 for the PSS and from .74 to .76 for SASS.

3.3.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. The starting point was the new typology totals for PSS and SASS. These are shown in table 3.2 below.

To carry out the Olkin GLS, the schools were placed into four school size classes (under 150, 150 to 499, 500 to 749, 750 and above).

After the Olkin adjustment, the difference between the PSS and SASS estimates had shrunk considerably (in overall absolute value) to

$$\underline{\mathbf{d}} = \begin{matrix} 19 \\ -125 \\ 0 \end{matrix}$$

The matrix M was obtained by tabulating the 1993-94 SASS file for the Catholic Private schools in the SASS sample. The values are

$$\begin{matrix} 144 & 5605 & 76274 \\ 5605 & 272841 & 3804568 \\ 76274 & 3804568 & 56973138 \end{matrix}$$

Solving for $\underline{\boldsymbol{\lambda}}$ yields

$$\underline{\boldsymbol{\lambda}}' = (+.79218, -0.02823, +0.00082)$$

and the Olkin GLS weights are of the form

$$u_i = w_i + 0.79218 - 0.02823t_i + 0.00082s_i$$

Notice that all the original weights are raised somewhat (by about .8); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered (usually they would not be lowered much except for large schools with many teachers). These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts are so small. One final comment, while the values for $\underline{\boldsymbol{\lambda}}$ are only shown to six significant digits, the calculations have been carried out in double precision.

3.3.3 Basic GLS Procedure Employed. -- To carry out the Basic GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 3.2 above that $\underline{\mathbf{d}}$ for the Basic GLS would be

$$\underline{\mathbf{d}} = \begin{matrix} 19 \\ -1169 \\ -17573 \end{matrix}$$

The matrix **M** is again obtained by tabulating the eligible portion of the 1993-94 SASS file for the Catholic Private schools in the SASS sample. Because the Olkin and Basic GLS employ just the unweighted sample to calculate **M**, it is the same for both (and hence not shown).

Solving for $\underline{\boldsymbol{\lambda}}$ yields this time

$$\underline{\boldsymbol{\lambda}}' = (+1.50665, -0.04082, +0.00040)$$

and the basic GLS weights are of the form

$$u_i = w_i + 1.50665 - 0.04082t_i + 0.00040s_i$$

Notice that again the original weights are raised this time (but by about twice the amount that the Olkin GLS weights were); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered. Usually they would not be lowered below what they were originally except for schools with large numbers of teachers and greater than average teacher/student ratios. These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts are so small. Looking just at the equation, concerns about negative weights arise but, as will be seen below, these did not materialize.

3.3.4 Operational Characteristics. -- To examine the Basic and Olkin GLS reweighting done, several "diagnostics" will be looked at. One statistic that may merit immediate attention is what happened to the spread in the weights themselves. Did the spread grow larger or smaller?

Figure 3.3 provides this information in its upper panels, which compare the original and two GLS adjustments. The Basic GLS weight has a larger spread than does the original SASS weight. This follows by noting that the original (or y) weights in the upper panel are slightly smaller than the GLS (or x) weights (since the equation which fits them is $y = .9452x$). For the Olkin GLS, the variability in the weights is greater still. The original SASS weights are related to the Olkin GLS weights by an equation of the form $y = .9184x$.

The overall differences in scale between the weights does not appear to be important. Still it is noticeable as the scatterplot of the Olkin GLS and original SASS weights shows.

The R^2 values in the upper panel of figures 3.3 might be commented on too. Despite the appearance of the scatter itself, very little evidence exists to indicate that either of the reweighting approaches altered the original weights in any major way. The R^2 values are both at or above 0.95 and most of the points lie just below the 45 degree line. The problem of negative weights did not arise for either the Olkin or Basic GLS methods. For the Basic GLS there was one school with a weight of less than one; for the Olkin GLS, there were two.

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 3.3 will continue to be our source. This time, though, look at the panel in the second row of the scatterplot matrix where the Basic and Olkin GLS are graphed. The Basic GLS (or y) values are slightly smaller than the Olkin GLS (or x) values; but ever so slightly. There is virtually no overall difference in the weights -- as evidenced by an R^2 of .97 between the two methods. The plotted points do indicate some departures though, as noted earlier; these are among the largest schools.

3.3.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Catholic Private Typology are available in tables 3.3 and 3.4, plus figure 3.4:

-- Table 3.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.

-- Table 3.4 is based on table 3.3 but focusses directly on percentage differences between the three SASS estimates and PSS.

-- Figure 3.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Notice that a log scale has been used here.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 3.3 and especially 3.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does the best on the average. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS 10/18 times. The Basic GLS tends to be closer to PSS than the original SASS (again, in 10 out of 18 comparisons). The results by community type are more mixed, as might be expected, since the Olkin approach did not try to control by community type (as it had by school size).

In figure 3.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already very close to the PSS, since the best fit regression equation which connects the various estimates is $y = .9941x$. There is just a little roughness around this average, however, as displayed visually and summarized by the R^2 value which is equal to $R^2 = .9839$. This is extremely good, suggesting that the SASS sample of Catholic Private schools is excellent.

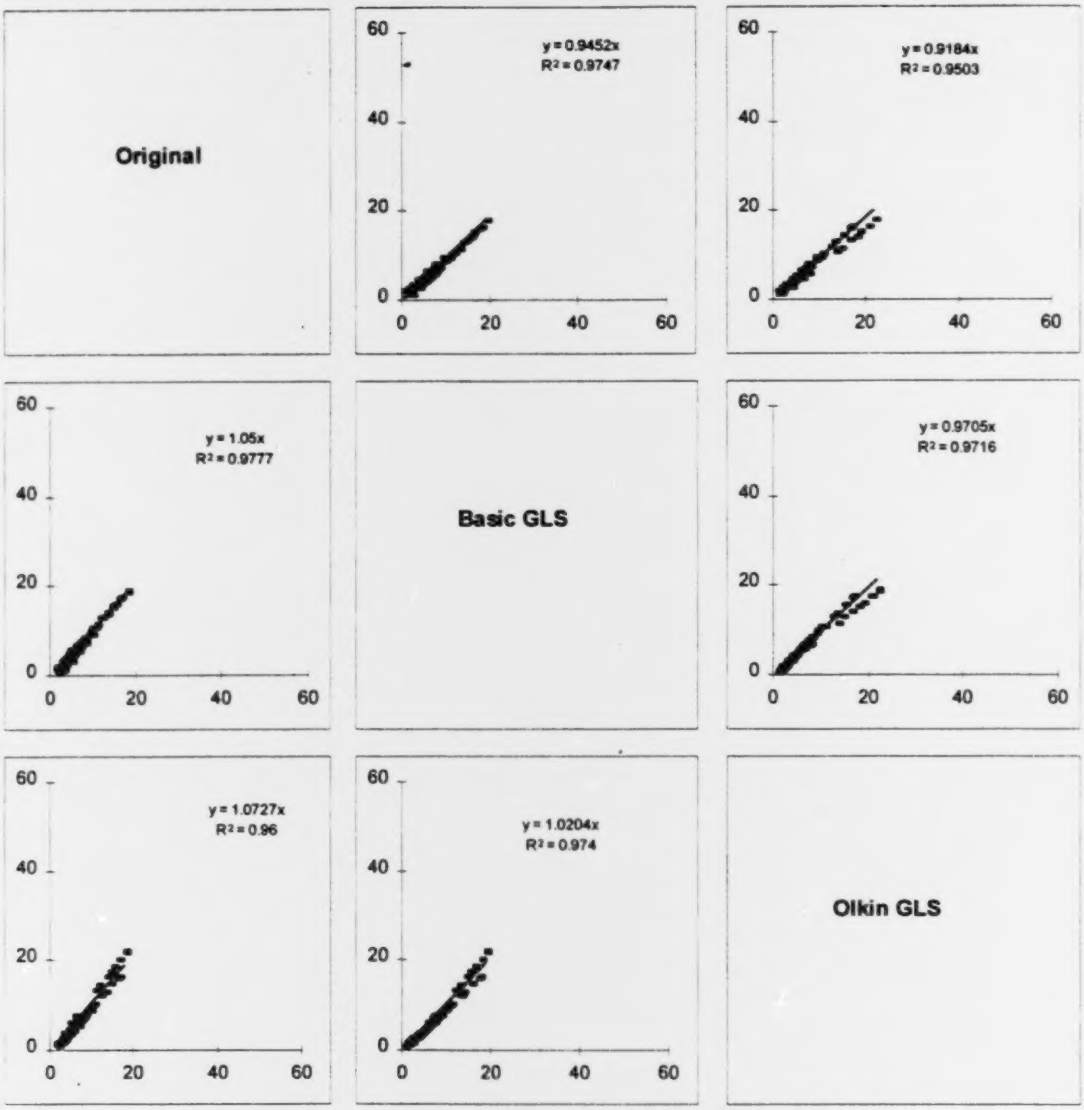
As in figure 3.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit is similar, yielding the relationship $y = .9952x$. Again, the average results for this method remain good. Somewhat less roughness is exhibited around the average as evidenced by the slightly larger R^2 value in this case ($R^2 = .9935$).

Finally, in figure 3.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields the relationship $y = .9935x$. The average results for this method are again comparable to the other two, with an R^2 value in this case of $R^2 = .9959$.

What can be concluded about this typology? Neither the Basic nor the Olkin GLS methods seem in any way inferior overall to the original SASS weighted file. To their credit, they both hit the overall PSS school, teacher, and student totals exactly. The Olkin method, furthermore and not surprisingly, does as well or better than the other two when estimates are looked at by school size.

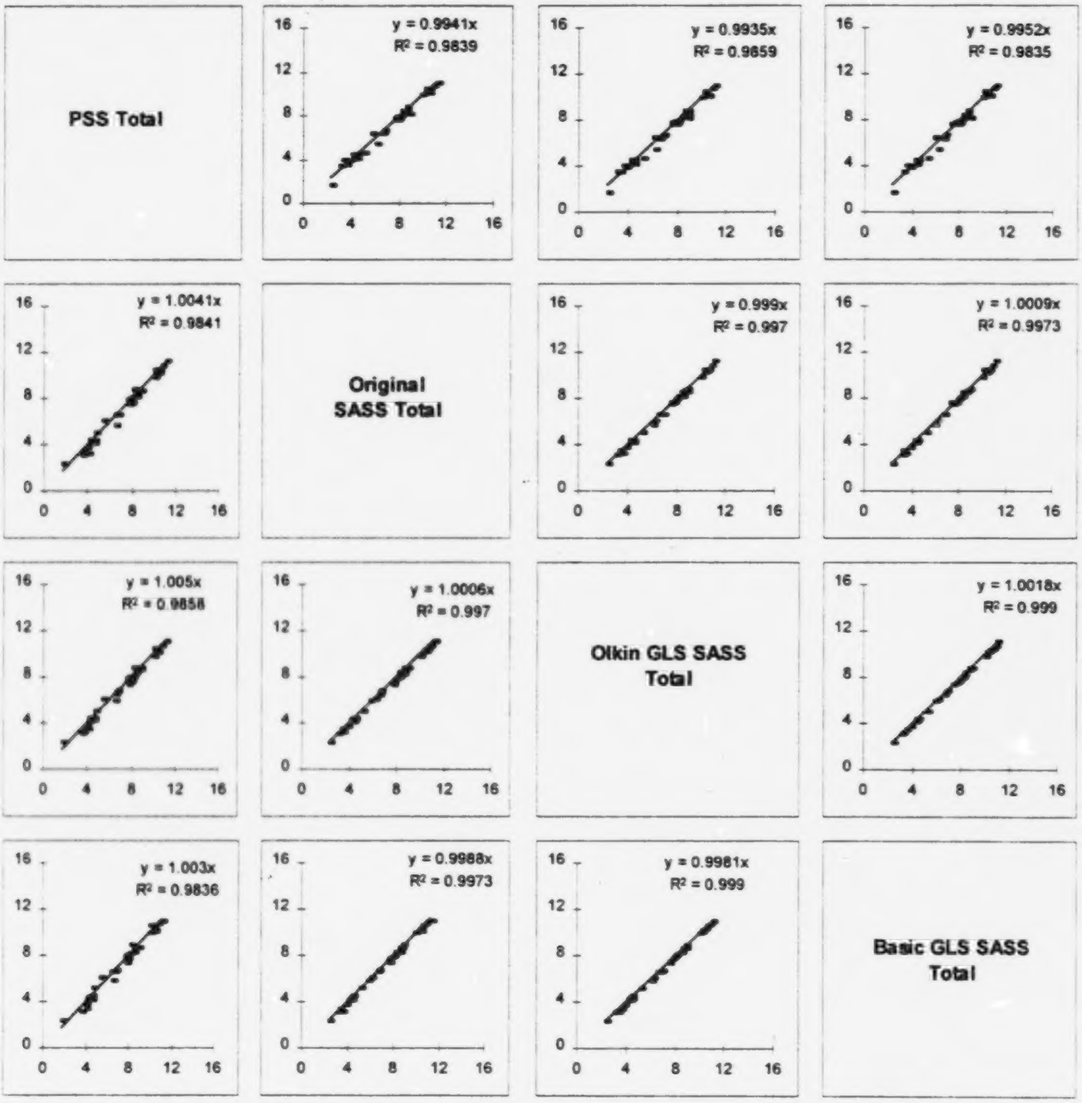
In the summary and recommendations section, some further comments will be made about how the Olkin GLS Approach might be improved further, leading to still better results.

Figure 3.3 -- Catholic Private: Scatterplot matrix comparing original, basic GLS, and Olkin GLS weights



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 3.4 – Catholic Private: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and basic SASS totals by school size and community type from Table 3.3 (in log scale)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 3.3 – Catholic Private: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
		Part I - PSS total (773 schools)			
1-149	School	88	60	52	200
	Teacher	803	705	606	2,114
	Student	6,520	4,429	4,443	15,392
150-299	School	100	90	31	222
	Teacher	1,977	2,063	584	4,625
	Student	22,127	20,280	6,625	49,032
300-499	School	88	65	18	171
	Teacher	2,624	2,273	558	5,455
	Student	34,506	26,496	6,847	67,849
500-749	School	74	48	5	127
	Teacher	3,118	2,213	221	5,552
	Student	44,651	29,137	3,244	77,032
750+	School	62	34	1	97
	Teacher	3,824	2,100	55	5,979
	Student	61,717	32,899	781	95,397
Total	School	413	297	107	817
	Teacher	12,346	9,355	2,024	23,724
	Student	169,522	113,241	21,940	304,702
		Part II - Original SASS total (144 schools)			
1-149	School	59	66	27	151
	Teacher	756	681	302	1,738
	Student	5,347	5,178	3,601	14,126
150-299	School	158	82	22	261
	Teacher	2,569	2,523	683	5,775
	Student	35,893	19,184	5,444	60,520
300-499	School	65	76	NA	141
	Teacher	2,043	2,844	NA	4,887
	Student	26,080	30,633	NA	56,713
500-749	School	80	43	10	133
	Teacher	3,331	1,853	434	5,618
	Student	48,184	26,739	6,837	81,760
750+	School	79	32	NA	111
	Teacher	4,764	2,111	NA	6,875
	Student	75,962	33,194	NA	109,155
Total	School	441	299	59	798
	Teacher	13,462	10,012	1,419	24,894
	Student	191,466	114,928	15,881	322,275

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 3.3 – Catholic Private: Estimates by school size and community type, PSS and SASS compared (cont'd)

Part III - Olkin GLS SASS total (144 schools)

School Size		Community Type: Central City	Community Type: Urban Fringe / Large Town	Community Type: Rural / Small Town	Total
1-149	School	73	82	34	189
	Teacher	939	846	382	2,167
	Student	6,638	6,436	4,535	17,609
150-299	School	162	81	21	264
	Teacher	2,614	2,478	664	5,756
	Student	36,758	19,030	5,344	61,132
300-499	School	67	76	NA	143
	Teacher	2,080	2,797	NA	4,877
	Student	26,744	30,581	NA	57,326
500-749	School	77	43	10	131
	Teacher	3,175	1,791	430	5,396
	Student	46,633	26,377	6,782	79,791
750+	School	65	26	NA	91
	Teacher	3,864	1,664	NA	5,528
	Student	62,483	26,361	NA	88,844
Total	School	445	307	65	817
	Teacher	12,672	9,576	1,476	23,724
	Student	179,256	108,785	16,661	304,702

Part IV - Basic GLS SASS total (144 schools)

School Size		Community Type: Central City	Community Type: Urban Fringe / Large Town	Community Type: Rural / Small Town	Total
1-149	School	64	71	30	165
	Teacher	820	738	340	1,898
	Student	5,787	5,601	3,999	15,386
150-299	School	173	87	23	282
	Teacher	2,792	2,630	698	6,120
	Student	39,248	20,278	5,631	65,156
300-499	School	71	79	NA	150
	Teacher	2,198	2,910	NA	5,107
	Student	28,308	32,011	NA	60,319
500-749	School	79	44	10	133
	Teacher	3,211	1,814	439	5,464
	Student	47,550	26,927	6,922	81,398
750+	School	63	24	NA	87
	Teacher	3,655	1,479	NA	5,135
	Student	59,271	23,171	NA	82,442
Total	School	450	305	63	817
	Teacher	12,677	9,570	1,477	23,724
	Student	180,164	107,987	16,551	304,702

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 3.4 -- Catholic Private: Estimates by school size and community type, PSS and SASS compared in percent difference

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and original SASS totals					
1-149	School	33.05	-9.68	48.62	24.31
	Teacher	5.92	3.47	50.19	17.78
	Student	18.00	-16.93	18.95	8.22
150-299	School	-57.31	9.82	30.15	-17.65
	Teacher	-29.94	-22.26	-16.94	-24.87
	Student	-62.21	5.41	17.83	-23.43
300-499	School	26.03	-16.35	NA	17.48
	Teacher	22.13	-25.13	NA	10.41
	Student	24.42	-15.61	NA	16.41
500-749	School	-8.23	9.59	-95.79	-5.11
	Teacher	-6.82	16.25	-96.60	-1.20
	Student	-7.91	8.23	-110.76	-6.14
750+	School	-26.80	4.57	NA	-14.54
	Teacher	-24.59	-0.53	NA	-14.99
	Student	-23.08	-0.90	NA	-14.42
Total %diff from PSS	School	-6.81	-0.47	45.05	2.29
	Teacher	-9.04	-7.03	29.88	-4.93
	Student	-12.94	-1.49	27.61	-5.77
Percent difference from PSS and Olkin GLS SASS totals					
1-149	School	16.93	-36.40	35.12	5.70
	Teacher	-16.89	-19.96	36.89	-2.51
	Student	-1.81	-45.33	-2.07	-14.41
150-299	School	-60.97	10.23	31.40	-18.96
	Teacher	-32.24	-20.10	-13.64	-24.47
	Student	-66.12	6.17	19.33	-24.68
300-499	School	24.12	-16.32	NA	16.50
	Teacher	20.72	-23.04	NA	10.60
	Student	22.49	-15.42	NA	15.51
500-749	School	-5.09	10.65	-94.26	-2.80
	Teacher	-1.81	19.07	-94.72	2.82
	Student	-4.44	9.47	-109.06	-3.58
750+	School	-4.80	23.79	NA	6.26
	Teacher	-1.06	20.75	NA	7.53
	Student	-1.24	19.87	NA	6.87
Total %diff from PSS	School	-7.68	-3.36	38.94	0.00
	Teacher	-2.64	-2.37	27.06	0.00
	Student	-5.74	3.93	24.06	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 3.4 – Catholic Private: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and Basic GLS SASS totals					
1-149	School	27.63	-18.83	42.55	17.60
	Teacher	-2.12	-4.57	43.85	10.23
	Student	11.24	-26.46	10.00	0.04
150-299	School	-71.91	4.14	27.66	-26.92
	Teacher	-41.25	-27.45	-19.51	-32.35
	Student	-77.37	0.01	15.01	-32.88
300-499	School	19.57	-21.95	NA	12.01
	Teacher	16.25	-28.02	NA	6.38
	Student	17.96	-20.81	NA	11.10
500-749	School	-7.47	8.59	-98.28	-5.14
	Teacher	-2.98	18.03	-98.75	1.58
	Student	-6.49	7.59	-113.38	-5.67
750+	School	-1.25	30.52	NA	10.89
	Teacher	4.40	29.55	NA	14.12
	Student	3.96	29.57	NA	13.58
Total %diff from PSS	School	-8.92	-2.47	41.25	0.00
	Teacher	-2.68	-2.31	27.01	0.00
	Student	-6.28	4.64	24.56	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

3.4 CONSERVATIVE CHRISTIAN TYPOLOGY

The Conservative Christian typology is a fairly large proportion of all private schools. For example, in the 1993-94 Private School Survey, there were an estimated 4,530 Conservative Christian schools or over 17% of the private school total for that year.

In table 4.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the SASS school total is larger than the PSS (by about 2.1%); SASS also estimates more teachers and students than are shown in PSS (0.5% and 4.9% more respectively).

To set the stage for the calculations that follow, it might be worth looking at figure 4.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black □'s. Overlaying the PSS data and shown as gray Δ's are the data from the 246 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS. The two slopes of the student/teacher relationship, though, are distinguishable, being 12.4 for PSS and 11.2 for SASS. While not directly comparable because of differences in sample designs, the R^2 values for the student/teacher ratios in both (unweighted) samples show analytically what can be observed from the graph, namely that the scatter is quite close around the average teacher/student relationship. These values are $R^2 = .83$ (PSS) and $R^2 = .80$ (SASS).

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing these new GLS estimators, a decision was first made about which sample cases to use (see section 3.4.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.4.2). The results of the basic GLS were also obtained (section 3.4.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and Basic GLS versions (section 3.4.4). An independent assessment (section 3.4.5) concludes the discussion.

3.4.1 Determining Outliers. -- Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 4.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

For the Conservative Christian typology, simple visual inspection resulted in reducing the PSS sample by 18 cases -- with a corresponding reduction in the SASS sample of 1 case.

Table 4.1 -- Conservative Christian: Weighted schools totals before excluding outliers (Based on 3,712 PSS and 246 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	4,530	4,626	-96
Teachers	51,289	51,562	-273
Students	610,578	640,369	-29,791

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

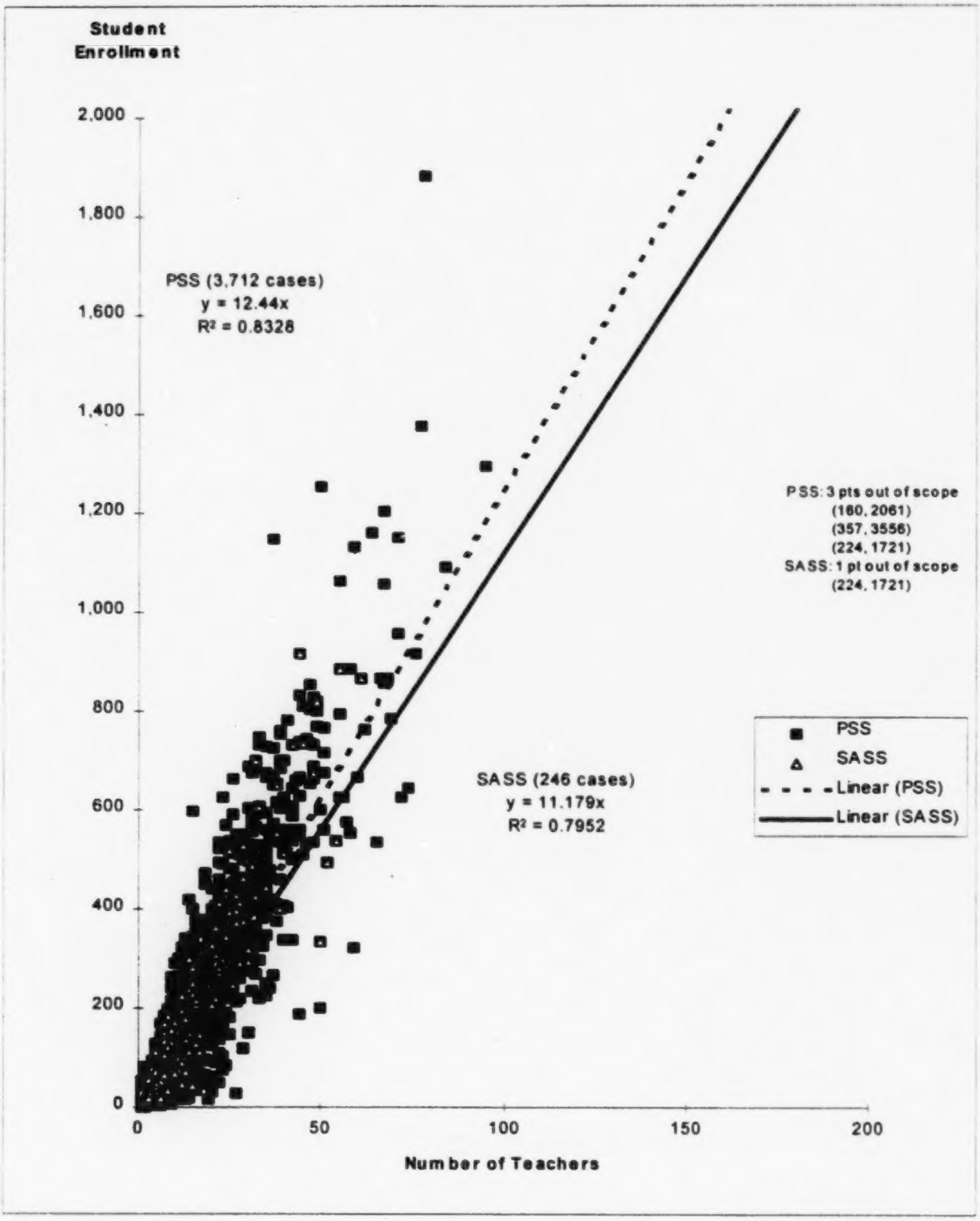
Table 4.2 -- Conservative Christian: Weighted schools totals after excluding outliers (Based on 3,690 PSS and 241 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	4,505	4,618	-113
Teachers	49,228	51,116	-1,888
Students	581,185	635,799	-54,613

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 4.1 – Conservative Christian: Student versus teacher unweighted sample total for PSS and SASS combined 1993-94

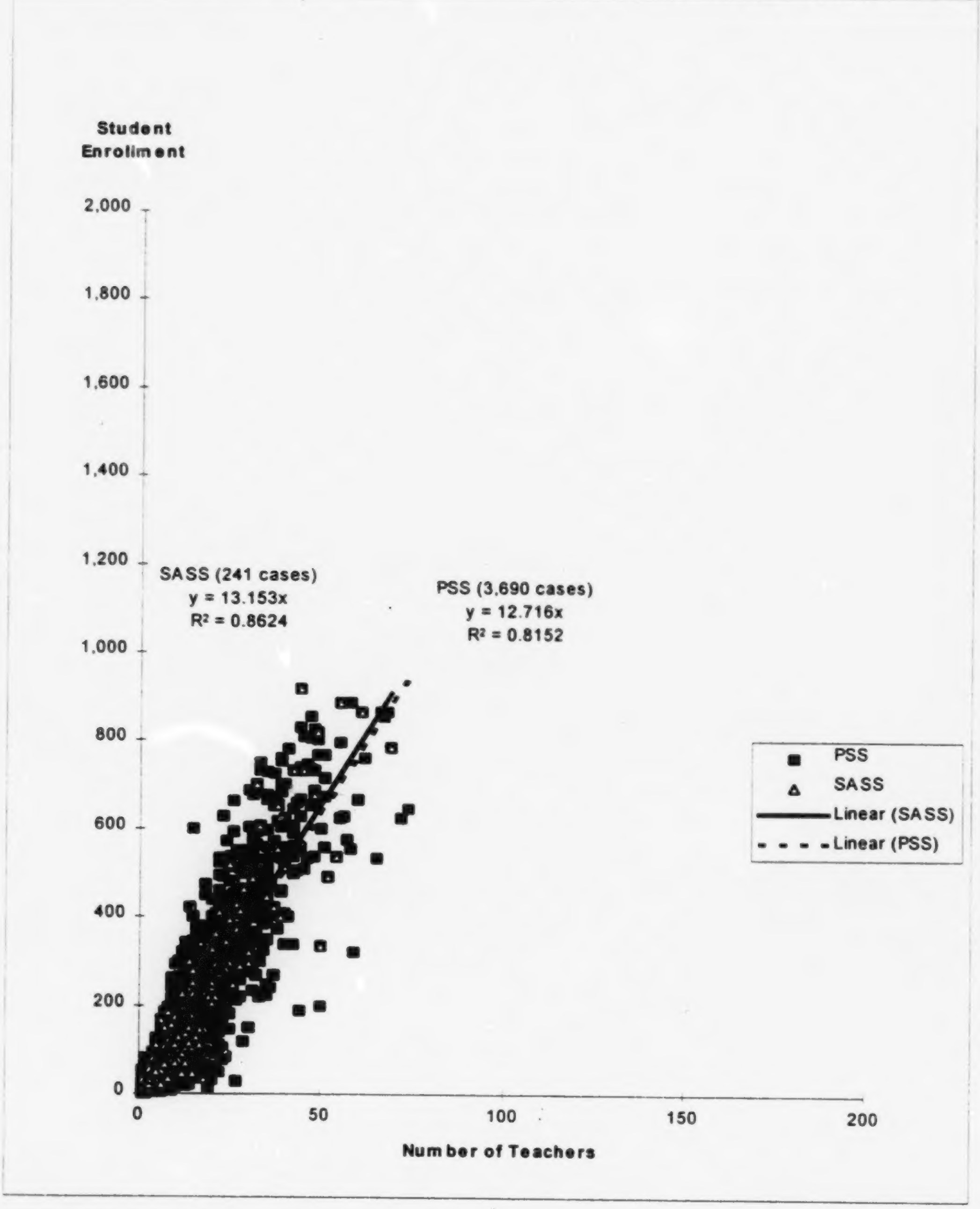
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 4.2 – Conservative Christian: Student versus teacher unweighted sample totals for PSS and SASS combined 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

When tried, this approach had to be supplemented by a more analytic method, which systematically excluded points more than a certain distance from the overall center of the combined PSS/SASS samples. After this second step, there was a further reduction of 4 PSS and 4 SASS points.

Figure 4.2 is the plot of the remaining 3690 PSS and 241 SASS cases. Notice that the student/teacher relationships have changed appreciably from those in figure 4.1. The student/teacher ratio in PSS went from 12.4 to 13.1; for SASS the ratio went from 11.2 to 12.7. These ratios now seemed close enough for the GLS method to have a chance of working without negative weights. Notice further, the scatter of points in both samples is visually even tighter. The R^2 values reflect this, rising from .77 to .80 for the PSS and from .74 to .76 for SASS.

3.4.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. The starting point was the new typology totals for PSS and SASS. These are shown in table 4.2 below.

To carry out the Olkin GLS, the schools were placed into three school size classes (under 150, 150 to 499, 500 and above).

After the Olkin adjustment, the difference between the PSS and SASS estimates had shrunk considerably (in overall absolute value) to

$$\underline{\mathbf{d}} = \begin{matrix} 89 \\ 1100 \\ -15515 \end{matrix}$$

The matrix **M** was obtained by tabulating the 1993-94 SASS file for the Conservative Christian schools in the SASS sample. The values are

$$\begin{matrix} 241 & 3457 & 43864 \\ 3457 & 83261 & 1095138 \\ 43864 & 1095138 & 15428854 \end{matrix}$$

Solving for $\underline{\boldsymbol{\lambda}}$ yields

$$\underline{\lambda}' = (-0.04124, +0.40067, -0.02933)$$

and the Olkin GLS weights are of the form

$$u_i = w_i - 0.04124 + 0.40067t_i - 0.02933s_i$$

Notice that all the original weights are lowered ever so slightly (by about .04); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered further (usually they would not be lowered, except for the schools with the very largest enrollments). Unlike for some of the other typologies, these additional school-by-school adjustments appear to be big -- given that the coefficients on the teacher and student counts are so large (Remember to put the teacher/student coefficients on a comparable basis, the student coefficient needs to be multiplied by roughly 13; while cancelling each other out near the center of the scatter, coefficients as large as shown should and did make for big changes in schools on either side. One final comment, while the values for $\underline{\lambda}$ are only shown to six significant digits, the calculations have been carried out in double precision.

3.4.3 Basic GLS Procedure Employed. -- To carry out the basic GLS weighting the equation

$$\underline{\lambda} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 4.2 above that $\underline{d}$ for the basic GLS would be

$$\underline{d} = \begin{matrix} -113 \\ -1888 \\ -54613 \end{matrix}$$

The matrix **M** is again obtained by tabulating the eligible portion of the 1993-94 SASS file for the Catholic Parochial schools in the SASS sample. Because the Olkin and Basic GLS employ just the unweighted sample to calculate **M**, it is the same for both (and hence not shown).

Solving for $\underline{\lambda}$ yields this time

$$\underline{\lambda}' = (-.85665, +0.41279, -0.03041)$$

and the basic GLS weights are of the form

$$u_i = w_i - 0.85665 + 0.41279t_i - 0.03041s_i$$

Notice that again the original weights are lowered, this time by quite a bit more than the amount that the Olkin GLS weights were; also, depending on the teacher and student counts in the sampled school, that school's weight may be increased or lowered further. These additional school-by-school adjustments are a concern -- given that the coefficients on the teacher and student counts are so large. In particular, concerns about negative weights arise; and, indeed, these did materialize.

3.4.4 Operational Characteristics. -- To examine the Basic and Olkin GLS reweighting done, several "diagnostics" will be looked at. One statistic that may merit immediate attention is what happened to the spread in the weights themselves. Did the spread grow larger or smaller?

Figure 4.3 provides this information in its upper panels, which compare the original and two GLS adjustments. The Basic GLS weight has a smaller spread than does the original SASS weight. This follows by noting that the original (or y) weights in the upper panel are slightly larger than the GLS (or x) weights (since the equation which fits them is $y = 1.0081x$). For the Olkin GLS, this continues to be true. The original SASS weights are related to the Olkin GLS weights by an equation of the form $y = 1.0169x$. As the scatterplot indicates, overall differences in scale between the weights does not appear to be important.

The R^2 values shown in the upper panel in figures 4.3 might be commented on too. Both are quite high, at or above 0.98 and most of the points lie very close to the 45 degree line. The problem of negative weights did not arise for the Olkin GLS method. There were, though, five schools with weights of less than one; for the Basic GLS, there were five negative weights and seven more less than one.

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 4.3 will continue to be our source. This time, though, look at the panel in the second row of the scatterplot matrix, where the Basic and Olkin GLS are graphed. The Basic GLS (or y) values are slightly larger than the Olkin GLS (or x) values; but ever so slightly. There is virtually no overall difference in the weights -- as evidenced by an R^2 of .997 between the two methods. The plot does indicate the problem noted earlier of small and negative weights.

3.4.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Conservative Christian Typology are available in tables 4.3 and 4.4, plus figure 4.4:

-- Table 4.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.

-- Table 4.4 is based on table 4.3 but focuses directly on percentage differences between the three SASS estimates and PSS.

-- Figure 4.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Notice that a log scale has been used here.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 4.3 and especially 4.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does the best on the average. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS 11 times. The Basic GLS also tends to be closer to PSS than the original SASS (again, 11 out of 18 times). The results by community type cannot be said to be very good for any of the estimators. Certainly the comparisons made in table 4.4 are mixed, as might be expected, since the Olkin approach did not try to control by community type (as it had by school size).

In figure 4.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already very close to the PSS, since the best fit regression equation which connects the various estimates is $y = .988x$. There is just a little roughness around this average, however, as displayed visually and summarized by the R^2 value which is equal to $R^2 = .9682$. This is extremely good, suggesting that, on the whole, the SASS sample of Conservative Christian schools is excellent.

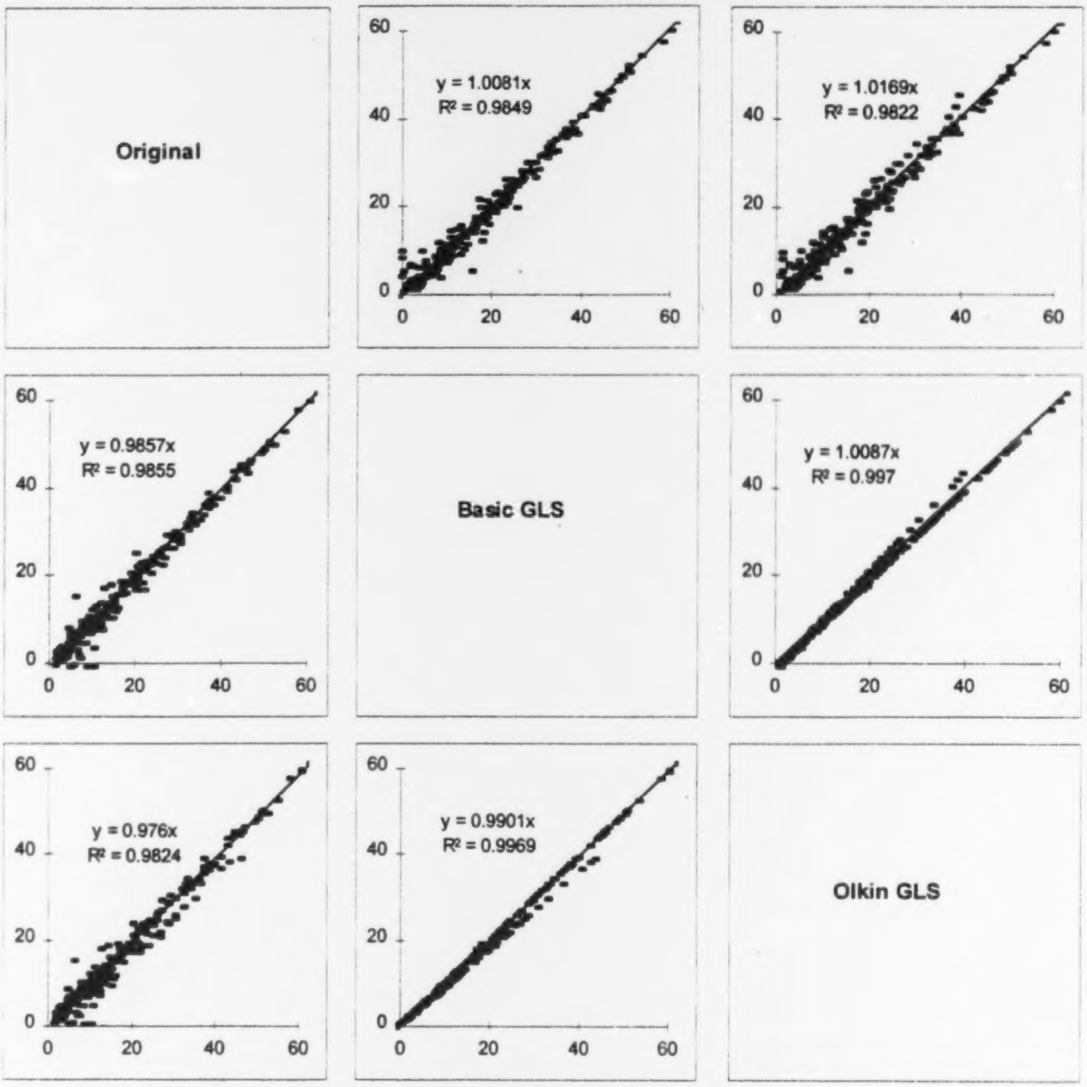
From figure 4.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit is similar, yielding the relationship $y = 1.0076x$. Again, the average results for this method remain good. A great deal more roughness is exhibited around the average, though, as evidenced by the much lower R^2 value in this case (with $R^2 = .9144$).

Finally, in figure 4.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields the relationship $y = 1.0029x$. The average results for this method are intermediate between the other two, with an R^2 value in this case of $R^2 = .9584$.

What can be concluded about this typology? The Olkin GLS method seems in no way inferior overall to the original SASS weighted file. To its credit, moreover, it hits the overall PSS school, teacher, and student totals exactly. The Olkin GLS method, furthermore and not surprisingly, does as well or better than the other two when estimates are looked at by school size. The Basic GLS method, while good in some respects, cannot be used without adjustment because of the negative weights which exist.

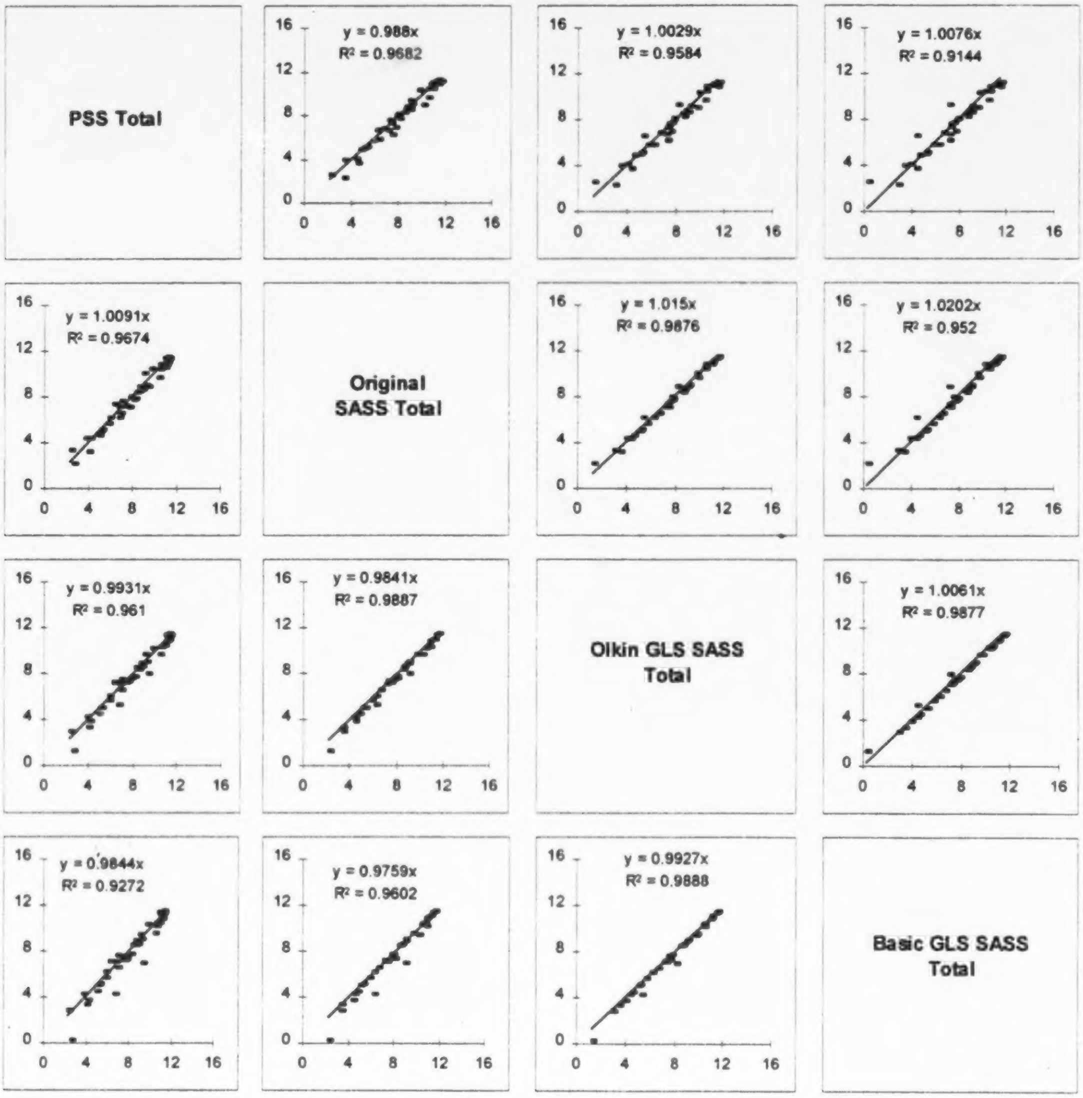
In the summary and recommendations section, additional comments will be made about how the Olkin might be improved further, leading to still better results.

Figure 4.3 – Conservative Christian: Scatterplot matrix comparing original, basic GLS, and Olkin GLS weights



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 4.4 – Conservative Christian: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and basic GLS SASS totals by school size and community type from Table 4.3 (in log scale)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94.

BEST COPY AVAILABLE

Table 4.3 -- Conservative Christian: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	

Part I - PSS total (3,690 schools)

School Size		Central City	Urban Fringe / Large Town	Rural / Small Town	Total
1-149	School	809	920	1,476	3,206
	Teacher	5,611	6,640	9,329	21,580
	Student	52,457	61,092	76,748	190,296
150-299	School	336	315	172	824
	Teacher	5,472	5,024	2,854	13,350
	Student	70,922	67,462	35,557	173,942
300-499	School	153	132	40	326
	Teacher	3,884	3,332	1,014	8,230
	Student	57,595	49,723	15,135	122,454
500-749	School	60	52	12	124
	Teacher	2,285	1,976	512	4,773
	Student	35,599	31,106	7,471	74,176
750+	School	10	13	2	25
	Teacher	497	725	73	1,295
	Student	8,160	10,959	1,199	20,317
Total	School	1,368	1,434	1,703	4,505
	Teacher	17,748	17,696	13,783	49,228
	Student	224,733	220,342	136,110	581,185

Part II - Original SASS total (241 schools)

School Size		Central City	Urban Fringe / Large Town	Rural / Small Town	Total
1-149	School	1,148	709	1,318	3,176
	Teacher	7,296	5,420	7,911	20,628
	Student	80,704	49,049	81,674	211,428
150-299	School	478	307	180	965
	Teacher	7,580	4,819	2,417	14,816
	Student	98,807	65,675	35,226	199,707
300-499	School	146	109	82	337
	Teacher	4,642	2,728	2,150	9,520
	Student	56,640	41,450	32,537	130,627
500-749	School	77	26	NA	104
	Teacher	2,808	1,258	NA	4,066
	Student	47,787	15,245	NA	63,032
750+	School	28	9	NA	37
	Teacher	1,624	462	NA	2,087
	Student	23,625	7,381	NA	31,006
Total	School	1,877	1,161	1,580	4,618
	Teacher	23,950	14,688	12,478	51,116
	Student	307,562	178,799	149,437	635,799

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 4.3 -- Conservative Christian: Estimates by school size and community type, PSS and SASS compared (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part III - Olkin GLS SASS total (241 schools)					
1-149	School	1,161	735	1,327	3,222
	Teacher	7,564	5,766	8,061	21,391
	Student	82,361	50,977	82,335	215,673
150-299	School	441	277	160	878
	Teacher	7,185	4,476	2,176	13,837
	Student	92,108	59,605	31,075	182,788
300-499	School	146	87	68	301
	Teacher	5,017	2,274	1,790	9,081
	Student	56,633	33,252	27,025	116,911
500-749	School	51	30	NA	81
	Teacher	1,909	1,474	NA	3,383
	Student	30,662	16,550	NA	47,213
750+	School	20	4	NA	23
	Teacher	1,343	192	NA	1,535
	Student	15,596	3,004	NA	18,601
Total	School	1,818	1,132	1,555	4,505
	Teacher	23,017	14,183	12,027	49,228
	Student	277,360	163,390	140,436	581,185
Part IV - Basic GLS SASS total (241 schools)					
1-149	School	1,141	719	1,312	3,172
	Teacher	7,390	5,628	7,952	20,971
	Student	80,582	49,660	81,292	211,534
150-299	School	470	292	169	931
	Teacher	7,621	4,692	2,284	14,598
	Student	97,995	62,640	32,794	193,428
300-499	School	153	90	74	316
	Teacher	5,237	2,348	1,932	9,518
	Student	59,220	34,382	29,178	122,780
500-749	School	42	27	NA	69
	Teacher	1,596	1,332	NA	2,928
	Student	25,210	14,685	NA	39,896
750+	School	16	1	NA	17
	Teacher	1,144	69	NA	1,213
	Student	12,438	1,110	NA	13,549
Total	School	1,822	1,128	1,555	4,505
	Teacher	22,989	14,070	12,169	49,228
	Student	275,445	162,477	143,263	581,185

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 4.4 -- Conservative Christian: Estimates by school size and community type, PSS and SASS compared in percent difference

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and original SASS totals					
1-149	School	-41.87	22.92	10.71	0.94
	Teacher	-30.04	18.37	15.20	4.41
	Student	-53.85	19.71	-6.42	-11.10
150-299	School	-42.02	2.59	-4.39	-17.07
	Teacher	-38.51	4.07	15.33	-10.98
	Student	-39.32	2.65	0.93	-14.81
300-499	School	4.91	17.48	-102.30	-3.30
	Teacher	-19.51	18.10	-112.01	-15.68
	Student	1.66	16.64	-114.97	-6.67
500-749	School	-29.74	49.33	NA	16.40
	Teacher	-22.90	36.33	NA	14.82
	Student	-34.24	50.99	NA	15.02
750+	School	-181.37	33.51	NA	-48.65
	Teacher	-227.11	36.24	NA	-61.12
	Student	-189.53	32.65	NA	-52.61
Total %diff from PSS	School	-37.16	19.01	7.22	-2.51
	Teacher	-34.94	17.00	9.47	-3.84
	Student	-36.86	18.85	-9.79	-9.40
Percent difference from PSS and Olkin GLS SASS totals					
1-149	School	-43.43	20.16	10.15	-0.51
	Teacher	-34.82	13.16	13.60	0.87
	Student	-57.01	16.56	-7.28	-13.34
150-299	School	-31.01	12.14	7.26	-6.49
	Teacher	-31.29	10.90	23.76	-3.65
	Student	-29.87	11.65	12.60	-5.09
300-499	School	4.80	34.46	-68.82	7.70
	Teacher	-29.17	31.73	-76.54	-10.35
	Student	1.67	33.13	-78.56	4.53
500-749	School	14.86	42.19	NA	34.82
	Teacher	16.46	25.37	NA	29.12
	Student	13.87	46.79	NA	36.35
750+	School	-98.99	73.09	NA	5.73
	Teacher	-170.48	73.50	NA	-18.54
	Student	-91.14	72.58	NA	8.45
Total %diff from PSS	School	-32.85	21.01	8.71	0.00
	Teacher	-29.69	19.85	12.74	0.00
	Student	-23.42	25.85	-3.18	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 4.4 – Conservative Christian: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		**Community Type**			**Total % diff from PSS**
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and Basic GLS SASS totals					
1-149	School	-41.01	21.91	11.14	1.07
	Teacher	-31.71	15.24	14.76	2.82
	Student	-53.61	18.71	-5.92	-11.16
150-299	School	-39.78	7.55	2.11	-12.91
	Teacher	-39.27	6.60	19.96	-9.35
	Student	-38.17	7.15	7.77	-11.20
300-499	School	0.33	32.25	-82.11	3.05
	Teacher	-34.84	29.51	-90.55	-15.65
	Student	-2.82	30.85	-92.78	-0.27
500-749	School	29.32	48.12	NA	44.25
	Teacher	30.14	32.59	NA	38.66
	Student	29.18	52.79	NA	46.21
750+	School	-62.34	89.96	NA	29.51
	Teacher	-130.42	90.44	NA	6.31
	Student	-52.44	89.87	NA	33.31
Total %diff from PSS	School	-33.17	21.29	8.73	0.00
	Teacher	-29.53	20.49	11.71	0.00
	Student	-22.57	26.26	-5.26	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

3.5 OTHER RELIGIOUS AFFILIATED TYPOLOGY

The Other Religious Affiliated typology is a fairly large proportion of all private schools. For example, in the 1993-94 Private School Survey, there were an estimated 3,640 schools or nearly 14% of the private school total for that year.

In table 5.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the SASS school total is smaller than the PSS (by about 4.1%); SASS also estimates less teachers and students than are shown in PSS (1.2% and 2.1% less respectively).

To set the stage for the calculations that follow, it might be worth looking at figure 5.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black **□'s.** Overlaying the PSS data and shown as gray **Δ's** are the data from the 575 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS. Nonetheless, the slopes of the two student/teacher relationship are clearly distinguishable, being 9.4 for PSS and 10.7 for SASS. While not directly comparable because of differences in sample designs, the R^2 values for the student/teacher ratios in both (unweighted) samples show analytically what can be observed from the graph, namely that most of the scatter is quite close around the average teacher/student relationship. These values are $R^2 = .77$ (PSS) and $R^2 = .76$ (SASS).

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing these new GLS estimators, a decision was first made about which sample cases to use (see section 3.5.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.5.2). The results of the basic GLS were also obtained (section 3.5.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and Basic GLS versions (section 3.5.4). An independent assessment (section 3.5.5) concludes the discussion.

3.5.1 Determining Outliers. -- Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 5.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

For the Other Religious Affiliated typology, simple visual inspection resulted in reducing the PSS sample by 32 cases -- with a corresponding reduction in the SASS sample of 10 case.

Table 5.1 -- Other Affiliated: Weighted schools totals before excluding outliers (Based on 3,176 PSS and 575 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	3,640	3,489	151
Teachers	52,237	51,612	625
Students	593,647	581,157	12,490

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

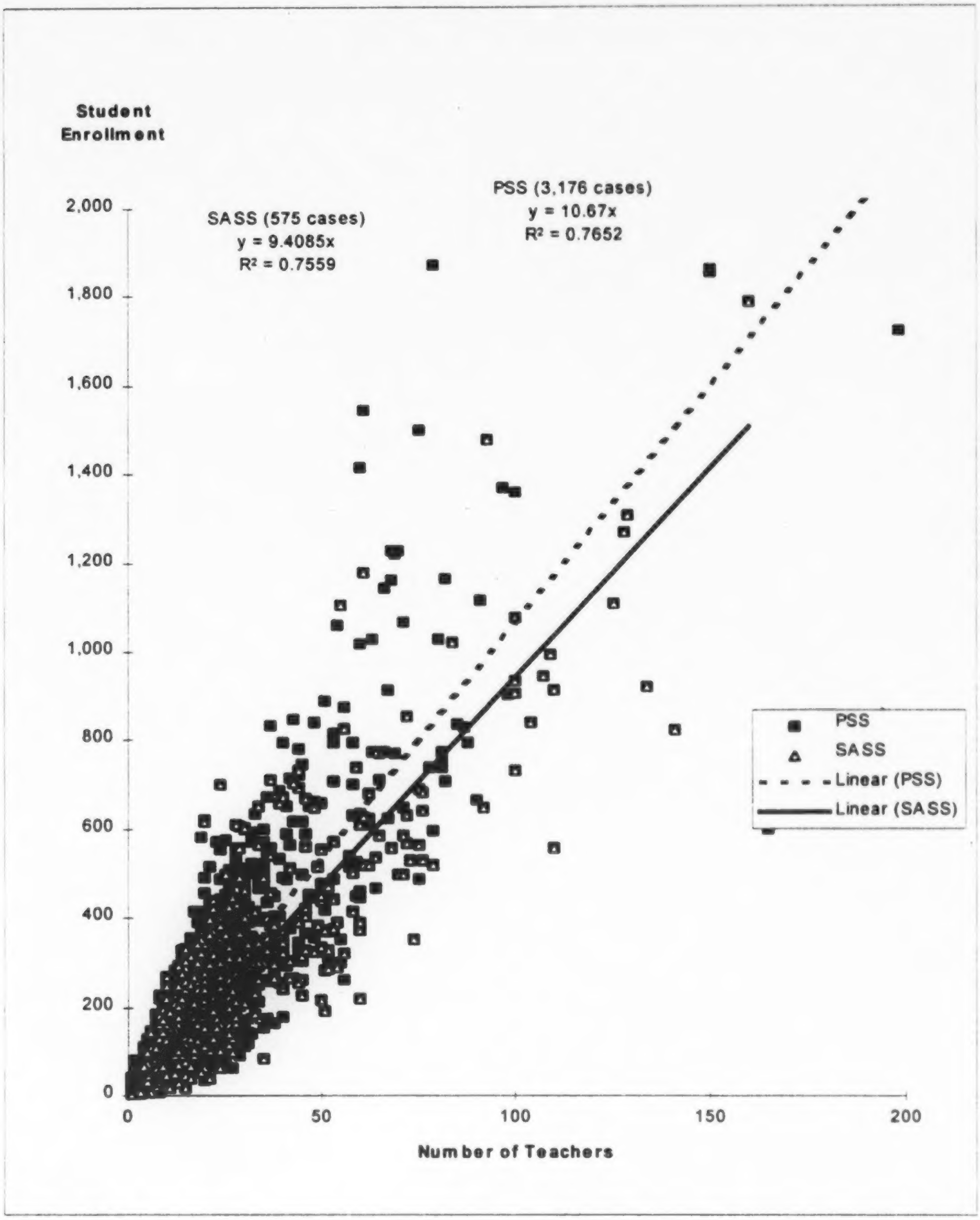
Table 5.2 -- Other Affiliated: Weighted schools totals after excluding outliers (Based on 3,144 PSS and 565 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	3,603	3,458	145
Teachers	48,674	48,329	346
Students	546,042	543,906	2,136

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 5.1 -- Other Affiliated: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

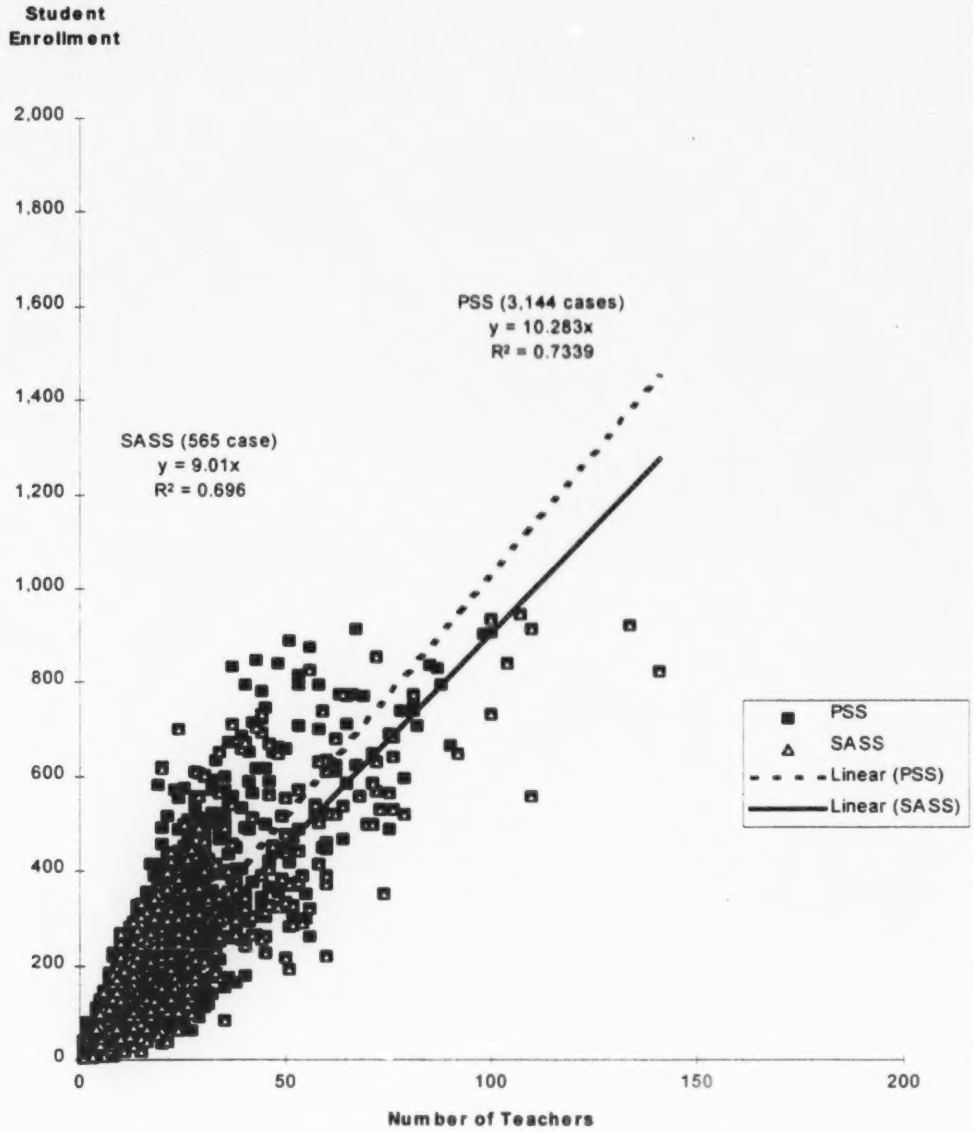
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94.

Figure 5.2 – Other Affiliated: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94.

BEST COPY AVAILABLE

This approach did not need to be supplemented by a more analytic method.

Figure 5.2 is the plot of the remaining 3144 PSS and 565 SASS cases. Notice that the student/teacher relationships have changed somewhat from those in figure 5.1. The student/teacher ratio in PSS went from 10.7 to 10.3; for SASS the ratio went from 9.4 to 9.0. Nonetheless, these seemed close enough for the GLS method to have a chance of working without negative weights. The R^2 values remain quite high but have dropped a little from 0.76 to 0.73 for the PSS and from 0.76 to 0.70 for SASS.

3.5.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. The starting point was the new typology totals for PSS and SASS. These are shown in table 5.2 below.

To carry out the Olkin GLS, the schools were placed into four school size classes (under 150, 150 to 499, 500 to 749, 750 and above).

After the Olkin adjustment, the difference between the PSS and SASS estimates had shrunk (considerably in overall absolute value) to

$$\underline{d} = \begin{matrix} -39 \\ -299 \\ 251 \end{matrix}$$

The matrix **M** was obtained by tabulating the 1993-94 SASS file for the Other Religious Affiliated schools in the SASS sample. The values are

$$\begin{matrix} 565 & 12686 & 127240 \\ 12686 & 530108 & 4776251 \\ 127240 & 4776251 & 49314734 \end{matrix}$$

Solving for $\underline{\lambda}$ yields

$$\underline{\lambda}' = (-0.16129, -0.00392, +0.0008)$$

and the Olkin GLS weights are of the form

$$u_i = w_i - 0.16129 - 0.00392t_i + 0.0008s_i$$

Notice that all the original weights are lowered somewhat (by about 0.2); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered further -- albeit slightly for the most part. Usually the weights would not be lowered, except for the schools with very large numbers of teachers. One final comment, while the values for $\underline{\boldsymbol{\lambda}}$ are only shown to six significant digits, the calculations have been carried out in double precision.

3.5.3 Basic GLS Procedure Employed. -- To carry out the basic GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 5.2 above that $\underline{\mathbf{d}}$ for the basic GLS would be

$$\underline{\mathbf{d}} = \begin{matrix} 145 \\ 346 \\ 2136 \end{matrix}$$

The matrix **M** is again obtained by tabulating the eligible portion of the 1993-94 SASS file for the Catholic Parochial schools in the SASS sample. Because the Olkin and Basic GLS employ just the unweighted sample to calculate **M**, it is the same for both (and hence not shown).

Solving for $\underline{\boldsymbol{\lambda}}$ yields this time

$$\underline{\boldsymbol{\lambda}}' = (.59302, -0.001139, -0.00137)$$

and the basic GLS weights are of the form

$$u_i = w_i + 0.59302 - 0.001139t_i - 0.00137s_i$$

Notice that this time the original weights are increased initially (by about .6, while for the Olkin GLS the intercept term was roughly -.2). Depending on the teacher and student counts in the SASS sampled school, the weights would be lowered. Looking just at the adjustment equation, it is unclear whether these additional school-by-school adjustments are a concern -- given that the coefficients on the teacher and student counts are so small. Negative weights could arise, of course; but these did not materialize. However, as with the Olkin GLS, there were weights less than one.

3.5.4 Operational Characteristics. -- Both the Basic and Olkin GLS reweighting done, as described above, seems to have worked well -- despite some of the resulting weights being on the small side. To indicate why this observation is made, several "diagnostics" will be looked at. One statistic that may merit immediate attention is what happened to the spread in the weights themselves. Did the spread grow larger or smaller?

Figure 5.3 provides this information in its upper panels, which compare the original and two GLS adjustments. The Basic GLS weight has a larger spread than does the original SASS weight. This follows by noting that the original (or y) weights in the upper panel are slightly smaller than the GLS (or x) weights (since the equation which fits them is $y = 0.9785x$). For the Olkin GLS, the variability in the weights is greater still; the original SASS weights are related to the Olkin GLS weights by an equation of the form $y = 0.9293x$.

The R^2 values shown in the upper panel in figures 4.3 might be commented on too. Both are quite high, at or above 0.99 and most of the points lie below but not too far from the 45 degree line. There were no problems with negative weights. However, for the Olkin GLS method, there were 68 schools with weights of less than one; similarly, for the Basic GLS, there were 22 schools with weights less than one. Incidentally, the schools with small weights did not have weights that were very much smaller than one, so no adjustment seemed needed.

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 5.3 will continue to be our source. This time, though, look at the panel in the second row of the scatterplot matrix where the Basic and Olkin GLS are graphed. Consistent with the story already told, the Basic GLS (or y) values are somewhat larger than the Olkin GLS (or x) values; but ever so slightly. The equation joining the two sets of weights is y (Basic GLS) = .9487x (Olkin GLS). There is virtually no overall distributional difference in the weights, beyond the difference in scale -- as evidenced by an R^2 of .994 between the two methods. The plot does indicate the problems noted earlier of some small weights.

3.5.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Other Affiliated Typology are available in tables 5.3 and 5.4, plus figure 5.4:

> -- Table 5.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.
>
> -- Table 5.4 is based on table 5.3 but focuses directly on percentage differences between the three SASS estimates and PSS.
>
> -- Figure 5.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Notice that a log scale has been used here.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 5.3 and especially 5.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does the best on the average. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS 14 times. The Basic GLS also tends to be closer to PSS than the original SASS (in 11 out of 18 comparisons). The results by community type are surprisingly good for both the GLS estimators. Of the two, the Olkin GLS is slightly to be preferred because it attempted to exercise some control by school size.

In figure 5.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already very close to the PSS, since the best fit regression equation which connects the various estimates is $y = 1.004x$. There is very little roughness around this average, too, as displayed visually and summarized by the R^2 value which is equal to $R^2 = .995$. This is extremely good, suggesting that the SASS sample of Other Affiliated schools is excellent.

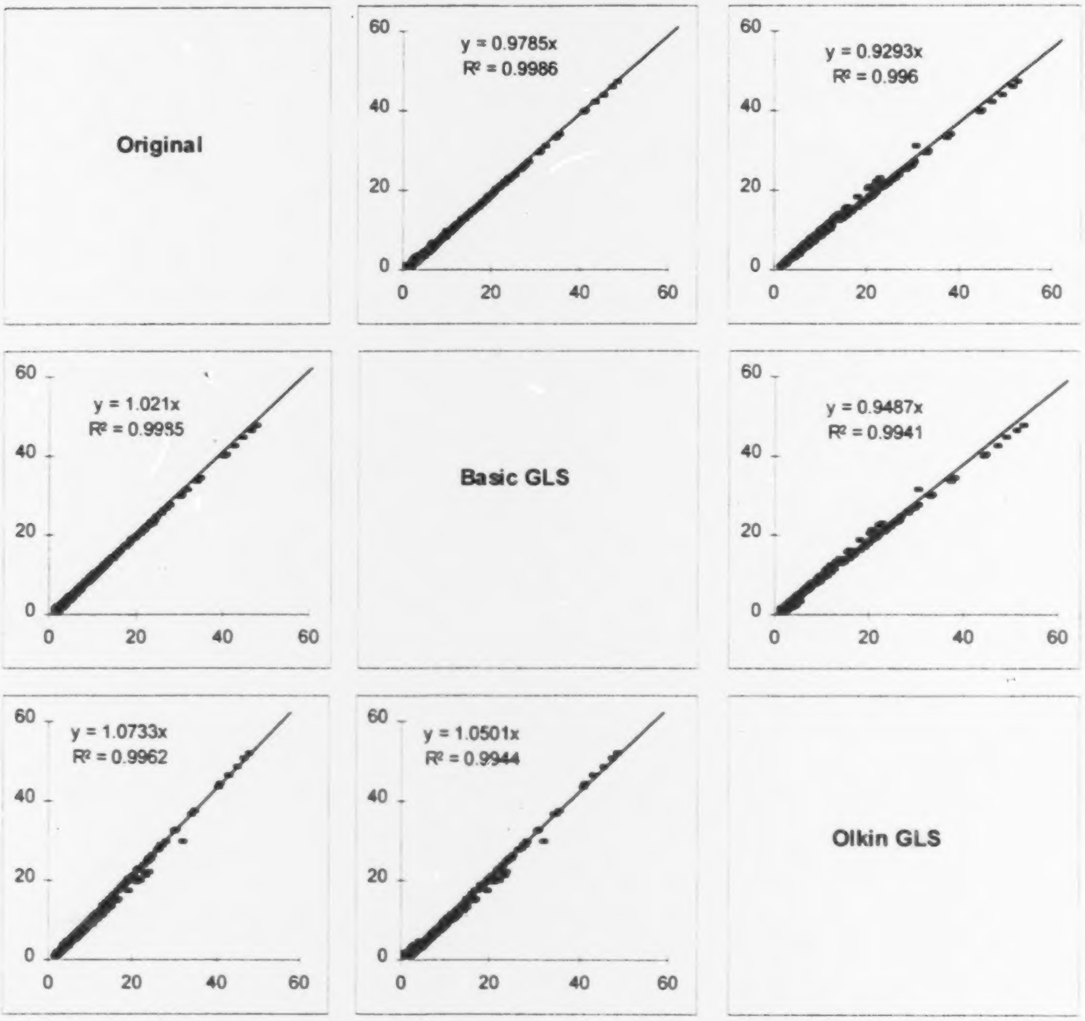
As in figure 5.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit is similar, yielding the relationship $y = 1.0077x$. Again, the average results for this method remain good. Slightly more roughness is exhibited around the average, as evidenced by a lower R^2 value in this case ($R^2 = .990$). Such a difference, obviously, is in no way important.

Finally, in figure 5.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields the relationship $y = 1.0016x$. This is the best of the three approaches, although all are close. For this estimator, the fit with the PSS is such that $R^2 = .9959$.

What can be concluded about this typology? The Olkin GLS method seems best overall. To its credit, moreover, it hits the PSS school, teacher, and student totals exactly. The Olkin method, furthermore and not surprisingly, does as well or better than the other two when estimates are looked at by school size. Some loss of sample efficiency arises because of the fact that some weights are slightly less than one but, while common, this seems to have no appreciable effects.

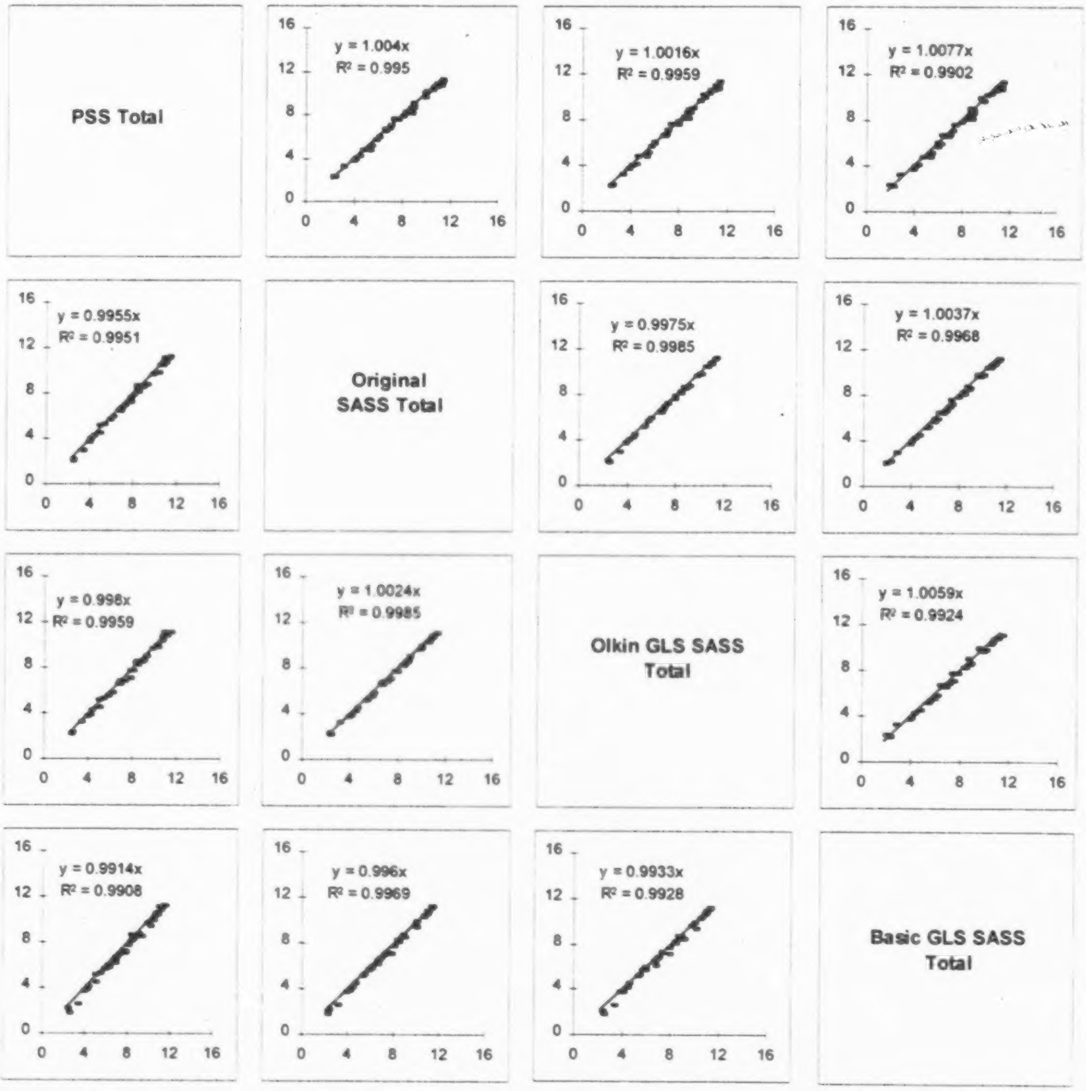
In the summary and recommendations section, some further comments will be made about how the Olkin might be improved further, leading to still better results.

Figure 5.3 -- Other Affiliated: Scatterplot matrix comparing original, basic GLS, and Olkin GLS weights



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94.

Figure 5.4 -- Other Affiliated: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and basic GLS SASS totals by school size and community type from Table 5.3
(in log scale)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94.

Table 5.3 – Other Affiliated: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type Central City	 Urban Fringe / Large Town	 Rural / Small Town	Total
Part I - PSS total (3,144 schools)					
1-149	School	766	715	834	2,315
	Teacher	5,952	5,401	3,938	15,290
	Student	55,009	52,033	38,601	145,643
150-299	School	380	284	121	784
	Teacher	7,084	4,991	1,949	14,023
	Student	80,819	61,016	25,393	167,228
300-499	School	181	119	42	343
	Teacher	5,658	3,580	1,150	10,387
	Student	66,749	45,222	15,418	127,389
500-749	School	63	53	9	125
	Teacher	3,256	2,614	413	6,284
	Student	38,159	31,950	5,592	75,702
750+	School	25	11	NA	36
	Teacher	1,918	772	NA	2,690
	Student	21,299	8,781	NA	30,080
Total	School	1,415	1,181	1,007	3,603
	Teacher	23,868	17,357	7,449	48,674
	Student	262,034	199,003	85,005	546,042
Part II - Original SASS total (565 schools)					
1-149	School	692	629	825	2,146
	Teacher	5,497	4,896	3,543	13,935
	Student	49,356	45,088	33,117	127,561
150-299	School	360	283	90	733
	Teacher	5,997	5,375	1,310	12,682
	Student	72,396	66,154	18,998	157,548
300-499	School	198	175	43	416
	Teacher	5,966	5,468	1,031	12,465
	Student	70,229	65,731	15,795	151,755
500-749	School	73	53	9	135
	Teacher	3,822	2,531	360	6,713
	Student	45,321	32,579	5,208	83,108
750+	School	20	8	NA	28
	Teacher	1,909	625	NA	2,534
	Student	17,369	6,565	NA	23,934
Total	School	1,343	1,148	967	3,458
	Teacher	23,191	18,894	6,243	48,329
	Student	254,671	216,117	73,118	543,906

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 5.3 – Other Affiliated: Estimates by school size and community type, PSS and SASS compared (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part III - Olkin GLS SASS total (565 schools)					
1-149	School	751	684	903	2,337
	Teacher	5,912	5,295	3,859	15,066
	Student	53,409	48,984	36,125	138,518
150-299	School	343	269	86	698
	Teacher	5,663	5,099	1,243	12,005
	Student	68,925	63,157	18,159	150,241
300-499	School	191	170	42	403
	Teacher	5,735	5,286	993	12,014
	Student	67,886	63,836	15,319	147,041
500-749	School	70	51	9	130
	Teacher	3,652	2,424	343	6,420
	Student	43,537	31,641	4,974	80,152
750+	School	25	10	NA	35
	Teacher	2,380	789	NA	3,169
	Student	21,818	8,273	NA	30,091
Total	School	1,380	1,184	1,039	3,603
	Teacher	23,343	18,893	6,438	48,674
	Student	255,574	215,891	74,576	546,042
Part IV - Basic GLS SASS total (565 schools)					
1-149	School	742	669	856	2,267
	Teacher	6,016	5,267	3,739	15,022
	Student	53,297	48,047	34,868	136,212
150-299	School	381	301	95	777
	Teacher	6,410	5,724	1,399	13,533
	Student	76,795	69,892	20,009	166,697
300-499	School	200	176	43	420
	Teacher	6,014	5,499	1,037	12,550
	Student	70,774	66,041	15,918	152,733
500-749	School	65	46	8	120
	Teacher	3,342	2,188	317	5,847
	Student	40,286	28,637	4,758	73,681
750+	School	14	6	NA	20
	Teacher	1,259	463	NA	1,722
	Student	11,773	4,947	NA	16,720
Total	School	1,402	1,199	1,003	3,603
	Teacher	23,041	19,141	6,492	48,674
	Student	252,924	217,564	75,553	546,042

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 5.4 -- Other Affiliated: Estimates by school size and community type, PSS and SASS compared in percent difference

Percent difference from PSS and original SASS totals

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
1-149	School	9.62	11.97	1.17	7.30
	Teacher	7.64	9.35	10.04	8.86
	Student	10.28	13.35	14.21	12.42
150-299	School	5.19	0.28	25.62	6.57
	Teacher	15.34	-7.69	32.79	9.57
	Student	10.42	-8.42	25.18	5.79
300-499	School	-9.37	-46.80	-1.90	-21.48
	Teacher	-5.45	-52.75	10.37	-20.00
	Student	-5.21	-45.35	-2.45	-19.13
500-749	School	-15.53	0.36	2.54	-7.47
	Teacher	-17.38	3.17	12.83	-6.84
	Student	-18.77	-1.97	6.87	-9.78
750+	School	20.93	24.22	NA	21.87
	Teacher	0.49	19.03	NA	5.81
	Student	18.45	25.24	NA	20.43
Total %diff from PSS	School	5.09	2.82	4.00	4.04
	Teacher	2.83	-8.86	16.19	0.71
	Student	2.81	-8.60	13.98	0.39

Percent difference from PSS and Olkin GLS SASS totals

School Size		Central City	Urban Fringe / Large Town	Rural / Small Town	Total % diff from PSS
1-149	School	2.02	4.31	-8.18	-0.95
	Teacher	0.67	1.95	2.01	1.47
	Student	2.91	5.86	6.42	4.89
150-299	School	9.69	5.01	28.95	10.97
	Teacher	20.05	-2.17	36.21	14.39
	Student	14.72	-3.51	28.49	10.16
300-499	School	-5.67	-42.38	1.23	-17.60
	Teacher	-1.36	-47.68	13.60	-15.66
	Student	-1.70	-41.16	0.65	-15.43
500-749	School	-10.79	3.50	7.09	-3.43
	Teacher	-12.17	7.27	16.93	-2.17
	Student	-14.09	0.97	11.06	-5.88
750+	School	0.67	4.56	NA	1.78
	Teacher	-24.11	-2.18	NA	-17.82
	Student	-2.44	5.79	NA	-0.04
Total %diff from PSS	School	2.50	-0.27	-3.18	0.00
	Teacher	2.20	-8.85	13.57	0.00
	Student	2.47	-8.49	12.27	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 5.4 – Other Affiliated: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	% diff from PSS
Percent difference from PSS and Basic GLS SASS totals					
1-149	School	3.12	6.40	-2.61	2.07
	Teacher	-1.07	2.47	5.05	1.76
	Student	3.11	7.66	9.67	6.48
150-299	School	-0.43	-6.02	21.55	0.94
	Teacher	9.51	-14.69	28.22	3.50
	Student	4.98	-14.55	21.20	0.32
300-499	School	-10.35	-47.89	-3.06	-22.52
	Teacher	-6.29	-53.61	9.81	-20.82
	Student	-6.03	-46.04	-3.24	-19.89
500-749	School	-2.95	12.30	10.69	4.51
	Teacher	-2.62	16.29	23.16	6.94
	Student	-5.57	10.37	14.92	2.67
750+	School	46.14	42.45	NA	45.05
	Teacher	34.36	39.97	NA	35.97
	Student	44.72	43.66	NA	44.41
Total %diff from PSS	School	0.95	-1.48	0.40	0.00
	Teacher	3.47	-10.28	12.85	0.00
	Student	3.48	-9.33	11.12	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

3.6 OTHER RELIGIOUS UNAFFILIATED TYPOLOGY

The Other Unaffiliated typology contains a fairly large fraction of all private schools. For example, in the 1993-94 Private School Survey, there were an estimated 4,051 such schools or over 15% of the private school total for that year.

In table 6.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the SASS school total is slightly smaller than the PSS (by about 0.7%); SASS also underestimates teachers (by 2.3%); but students are overestimated in SASS relative to PSS (by 8.8%).

To set the stage for the calculations that follow, it might be worth looking at figure 6.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black **□'s**. Overlaying the PSS data and shown as gray **Δ's** are the data from the 148 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS. Nonetheless, the slopes of the student/teacher relationship are different, being 9.9 for PSS and 10.7 for SASS. While not directly comparable because of differences in sample designs, the R^2 values for the student/teacher relationships in both (unweighted) samples are of moderate size, showing analytically what can be observed from the graph. These values are $R^2 = .85$ (PSS) and $R^2 = .73$ (SASS). The presence of some very large influential observations, not shown because of their size, is also a factor.

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing these new GLS estimators, a decision was first made about which sample cases to use (see section 3.6.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.6.2). The results of the Basic GLS were also obtained (section 3.6.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and Basic GLS versions (section 3.6.4). An independent assessment (section 3.6.5) concludes the discussion.

3.6.1 Determining Outliers. --Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 6.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

Table 6.1 -- Other Unaffiliated: Weighted schools totals before excluding outliers (Based on 3,193 PSS and 329 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	4,051	4,023	28
Teachers	38,410	37,515	895
Students	425,356	462,934	-37578

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

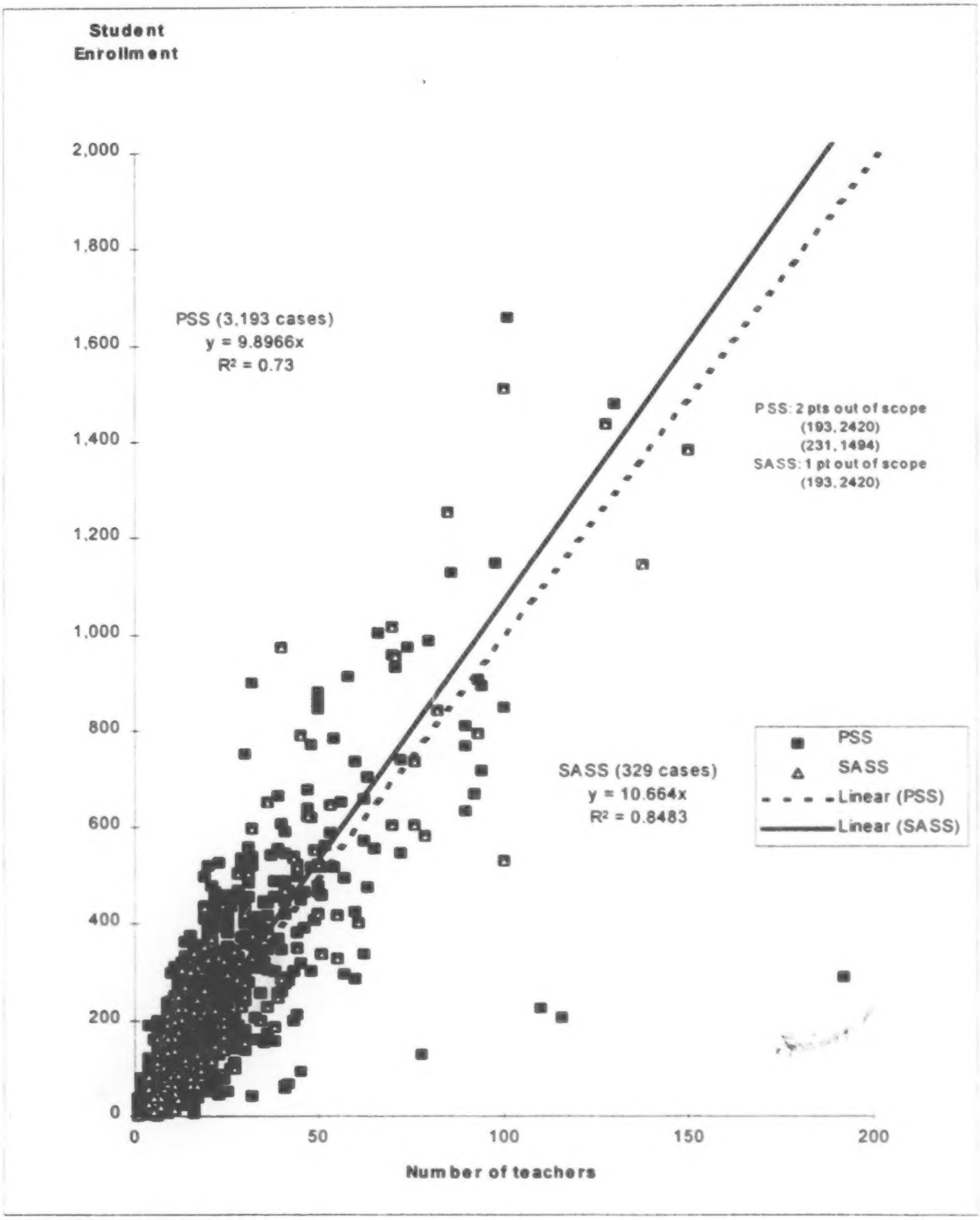
Table 6.2 -- Other Unaffiliated: Weighted schools totals after excluding outliers (Based on 3,141 PSS and 313 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	3,994	3,675	319
Teachers	33,523	29,401	4,122
Students	373,168	345,480	27,688

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 6.1 – Other Unaffiliated: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

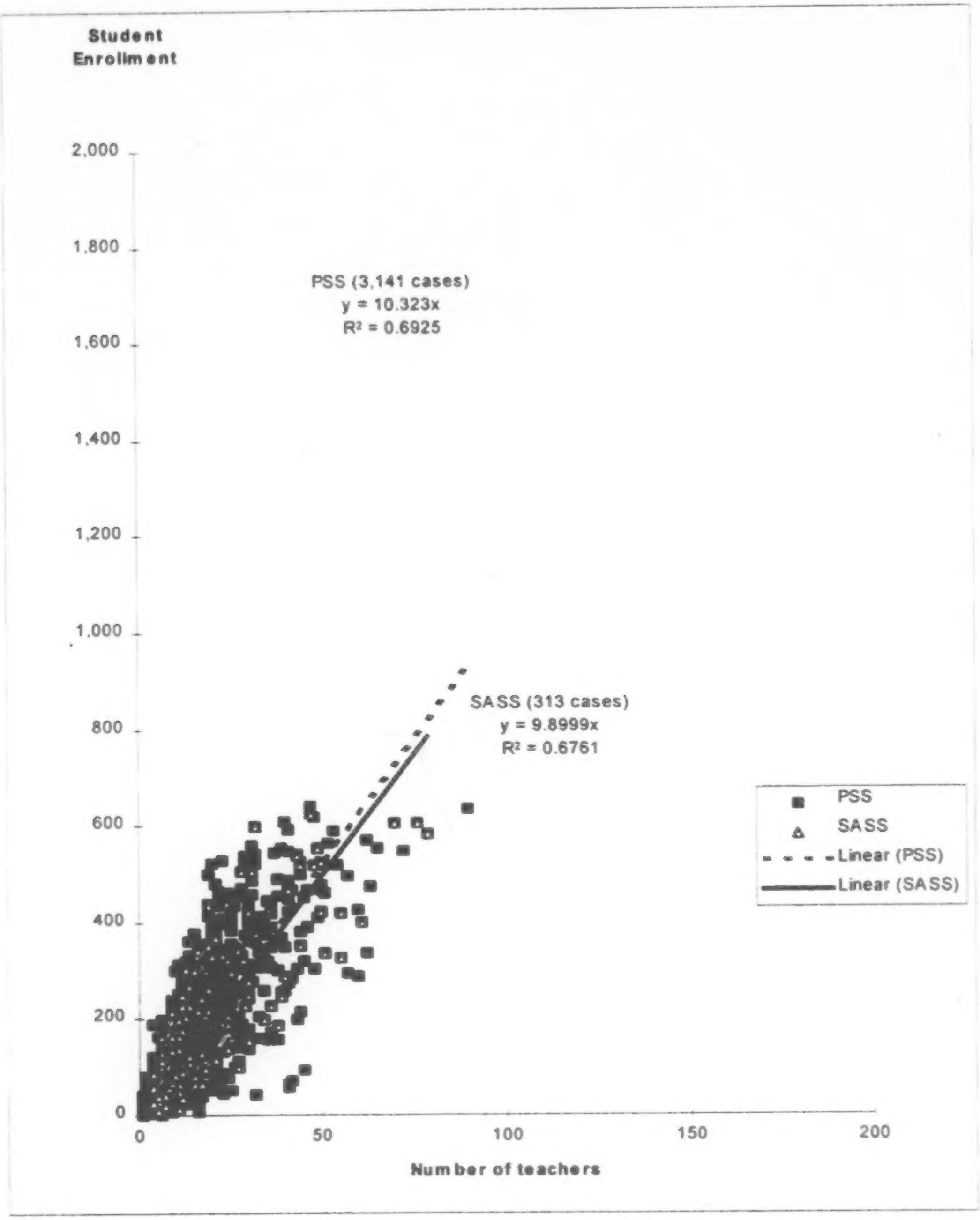
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94.

Figure 6.2 – Other Unaffiliated Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94.

For the Other Affiliated typology, simple visual inspection resulted in reducing the PSS sample by 49 cases -- with a corresponding reduction in the SASS sample of 14 cases.

A series of trial and error steps, then, took place to locate still other outliers. Eventually the data had to be grouped into six separate school sizes and GLS run separately on each. This lead to the elimination of 3 more PSS schools and 2 more SASS cases.

Figure 6.2 is the plot of the remaining 3141 PSS and 313 SASS cases. Notice that the student/teacher relationships are little changed overall from those in figure 6.1; however, the scatter of points in both samples is visually a lot tighter. Ironically, though, because of the influential observations eliminated, the R^2 values actually decline, falling from .73 to .69 for the PSS and from .85 to .68 for SASS.

3.6.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\lambda} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. Overall, these new typology totals for PSS and SASS are shown in table 6.2 below.

Unlike with the other typologies, to carry out the Olkin GLS, the schools were placed into six school size classes (specifically, under 66, 66 to 149, 150-199, 200-249, 250-399, and 400+). The SASS samples in these categories, ranged from 19 to 102. This may have been too many groups. Certainly a group with only 19 observations was borderline.

To provide a feel for the difficulties encountered, figure 6.5 might be examined. It compares the PSS and original SASS cumulative weight distributions by student enrollment within each group separately. From it, we can see widely divergent patterns -- from near total agreement in the distributions in some groups (e.g., the under 66 group), to groups where the PSS and SASS bear almost no relationship to one other (e.g., the 200 to 399 group). These differences in distribution made the typology extremely hard to handle; in particular, it was quite hard to avoid negative weights.

For the Olkin GLS, what was done, unlike elsewhere, was to directly (and successfully) attempt a GLS adjustment within each of the six categories. An initial step was to truncate or bound the original SASS weight distributions within each class so that their variability was reduced. Separate bounds were used to reflect the downward shift of the average weight as the schools increased in enrollment size. A scaling factor , the Olkin Adjustment, was introduced next, as described in Section 2.2. Because our adjustments were being done in classes similar to those used in the original sample design, arguably they should not have affected the efficiency of the SASS sample greatly; moreover, our approach directly attacked the variability of the original weights which was considerable.

Rather than spell out all the adjustments in detail, it may be enough to compare the cumulative Olkin/PSS weight plots for each group separately at each step. Figure 6.5, which has already been discussed, provides the starting point, since it displays the original SASS weight distributions and compares them with PSS. In figure 6.6, the effect of the truncation step is shown. Notice how this step, all by itself, improved the appearance of most of the comparisons. In figure 6.7, the final Olkin GLS adjustment completes the job. More on this approach and related alternatives is found in Section 4, under recommendations for future study.

3.6.3 Basic GLS Procedure Employed. -- To carry out the basic GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 6.2 above that **d** for the basic GLS would be

$$\underline{\mathbf{d}} = \begin{matrix} 57 \\ 2298 \\ -9620 \end{matrix}$$

The overall matrix **M** is obtained by tabulating the eligible portion of the 1993-94 SASS file for the entire set of Other Unaffiliated schools in the SASS sample. **M** is

$$\begin{matrix} 313 & 3908 & 44421 \\ 3908 & 100010 & 990091 \\ 44421 & 990091 & 11477607 \end{matrix}$$

Solving for $\underline{\boldsymbol{\lambda}}$ yields this time

$$\underline{\boldsymbol{\lambda}}' = (+.55248, +0.21135, -0.02121)$$

and the basic GLS weights are of the form

$$u_i = w_i + 0.55248 + 0.21135t_i - 0.02121s_i$$

Notice that the original weights are raised (by about .6); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered. These additional school-by-school adjustments appear modest. Still, just looking at the equation, there are concerns about negative weights; and, as will be seen below, these do materialize.

3.6.4 Operational Characteristics. -- Figure 6.3 provides this information on operational characteristics in its upper panels; these compare the original and two GLS adjustments. The Basic GLS weight has a slightly larger overall spread than does the original SASS weight. This follows by noting that the original (or y) weights in the upper panel are slightly smaller than the GLS (or x) weights (since the equation which fits them is y = 0.993x). Now, on the other hand, the original SASS weights are related to the Olkin GLS weights by an equation of the form y = 1.1234x. Hence, the spread in the Olkin (or x) weights is considerably smaller than for the original (or y) weights.

The overall differences in scale between the weights does not appear to be important for the Basic GLS and original SASS comparisons. This is not the case for the Olkin GLS where the scatter is quite disturbed.

The R^2 values shown in the upper panel in figures 6.3 might be commented on too. For the Basic GLS results, the R^2 values is high above 0.99 and most of the points lie just below the 45 degree line. For the Olkin GLS comparison with the original SASS weights, the R^2 value is less satisfactory at $R^2 = .84$. The problem of negative weights is evident in the graph and did occur for the Basic GLS, with 9 negatives and an additional 19 cases less than one. For the Olkin GLS, by design, there were no negative weights but there were 10 cases somewhat less than one.

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 6.3 will continue to be our source. This time, though, look at the panel in the second row of the scatterplot matrix where the Basic and Olkin GLS are graphed. The Basic GLS (or y) values are slightly smaller than the Olkin GLS (or x) values; but ever so slightly. The pattern of overall difference observed for the original SASS weights continues -- as evidenced by an R^2 of just over .84 between the two methods. The plotted points do indicate numerous departures, especially among the larger schools.

3.6.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Other Unaffiliated Typology are available in tables 6.3 and 6.4, plus figure 6.4:

-- Table 6.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.

-- Table 6.4 is based on table 6.3 but focuses directly on percentage differences between the three SASS estimates and PSS.

-- Figure 6.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Note this figure has been plotted in log scale.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 6.3 and especially 6.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does the best on the average. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS 14/18 times. It "betters" the Basic GLS in closeness to PSS (being closer in 15 out of 18 comparisons). The results by community type are also encouragingly favorable to the Olkin approach, even though there were no controls by community type (as had been true by school size).

In figure 6.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already very close to the PSS, since the best fit regression equation which connects the various estimates is $y = 1.0029x$. There is some roughness around this average, however, as displayed visually and summarized by the R^2 value which is equal to $R^2 = .9504$.

As in figure 6.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit is similar, yielding the relationship $y = 1.0x$. Again, the average results for this method remain good. Somewhat less roughness is exhibited around the average as evidenced by the larger R^2 value in this case ($R^2 = .9839$).

Finally, in figure 6.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields the relationship $y = 1.0015x$. The average results for this method are again comparable to the other two, although the R^2 value in this case is the smallest, at $R^2 = .9372$.

What can be concluded about this typology? Neither the Basic nor the Olkin GLS methods seem in any way inferior overall to the original SASS weighted file. To their credit, they both hit the overall PSS school, teacher, and student totals exactly. The Olkin GLS method might be preferred, because it has no negative weights (unlike the Basic GLS).

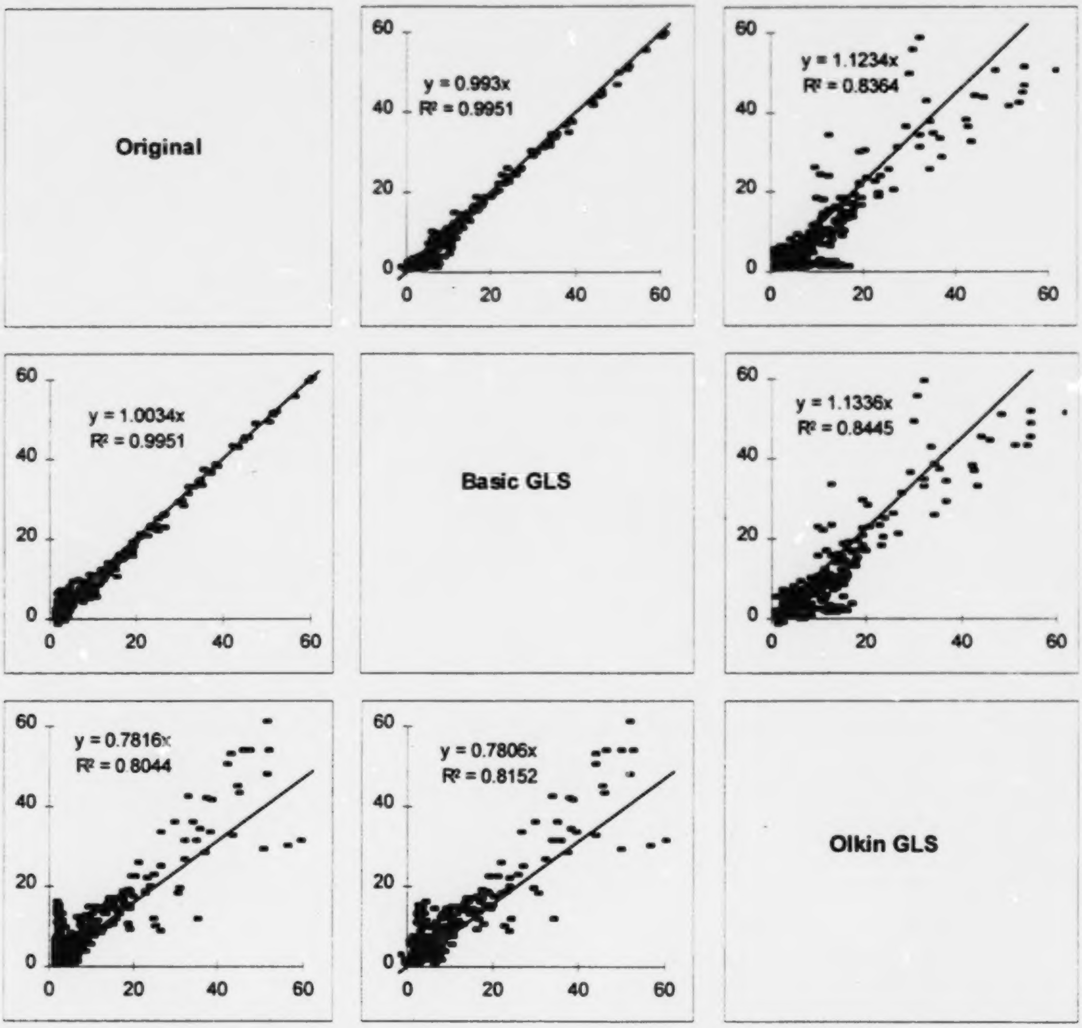
In the summary and recommendations section, some further comments will be made about how the Olkin GLS might be improved further, leading to still better results.

3.6.6 Cumulative Weight Plots. -- Figures 6.5 to 6.7 track the Olkin GLS approach taken for the other unaffiliated typology. First in figure 6.5 the original PSS and SASS cumulative weight distributions are compared by number of students for each of the 6 groups eventually adjusted separately. Notice that for the smaller schools the PSS and SASS are not too far apart. However large departures occur among the larger schools.

Figure 6.6 shows what happens to the cumulative weight distributions if the original SASS weights are simply bounded slightly. Evident improvements occur but gaps between SASS and PSS also clearly exist.

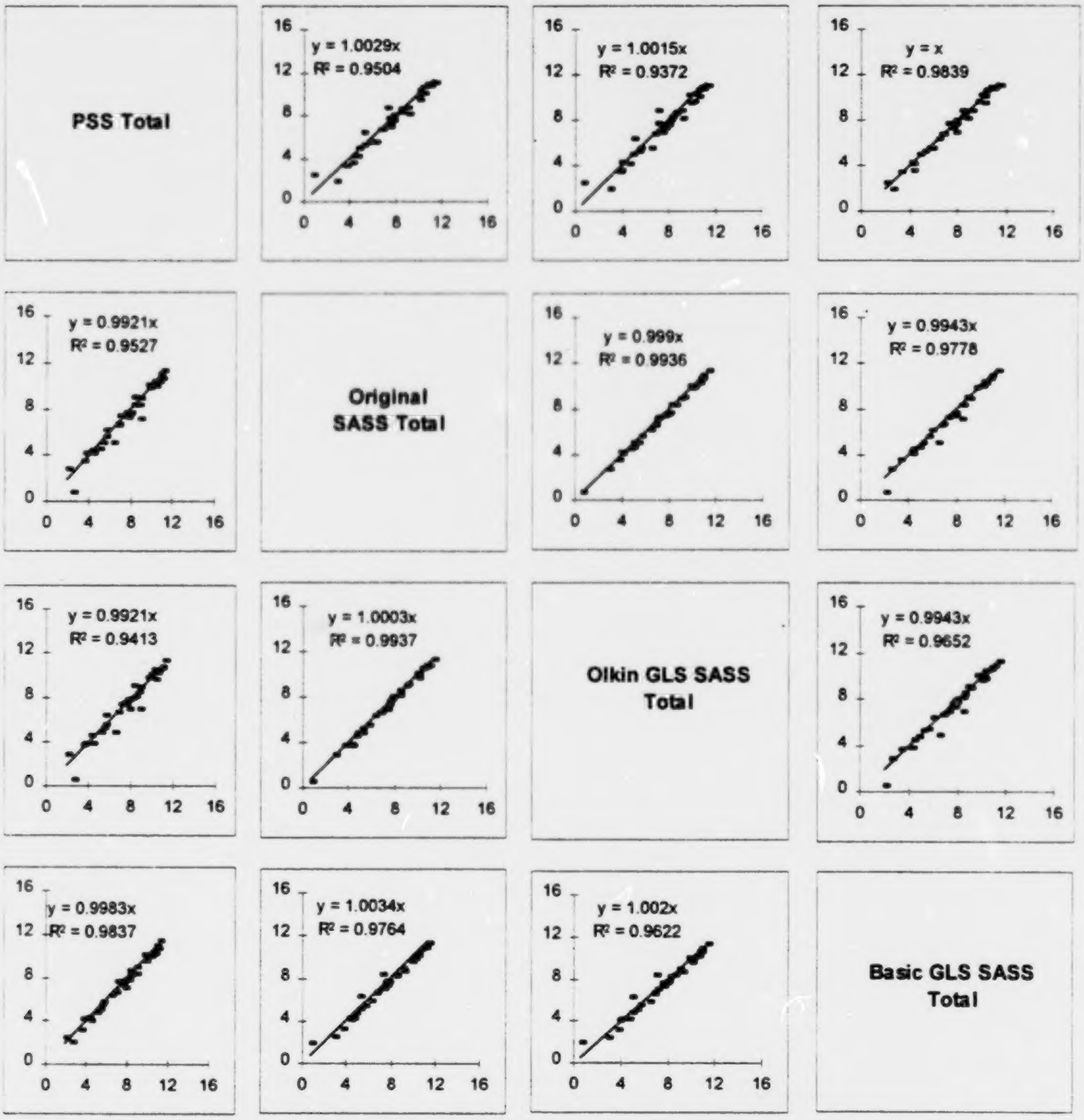
Figure 6.7 shows what the final Olkin GLS weights achieved. A very regular pattern!

Figure 6.3 -- Other Unaffiliated: Scatterplot matrix comparing original, basic GLS, and Olkin GLS weights



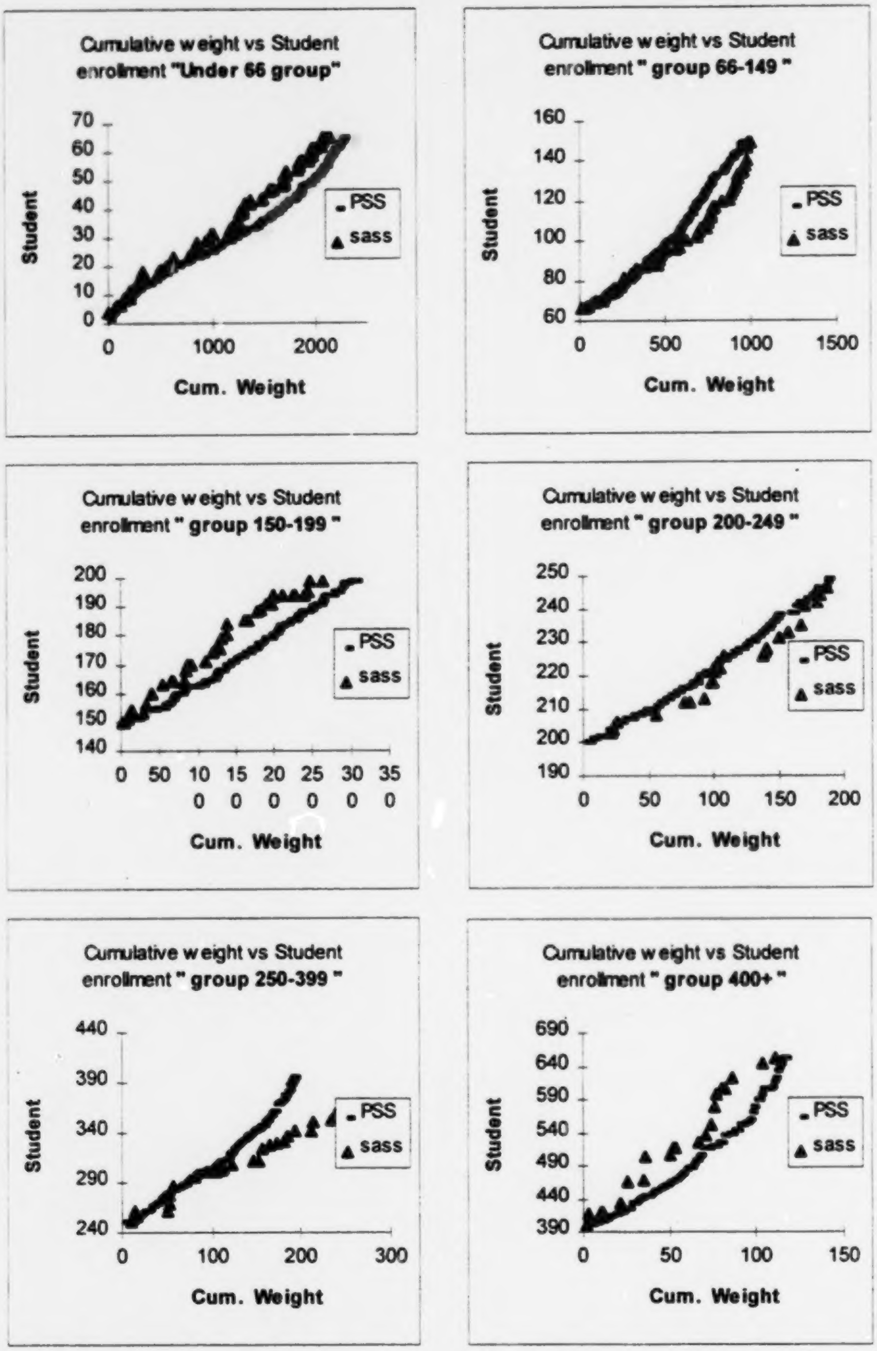
SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Surveys, 1993-94, Private School Surveys, 1993-94.

Figure 6.4 -- Other Unaffiliated: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and basic GLS SASS totals by school size and community type from Table 6.3 (in log scale)



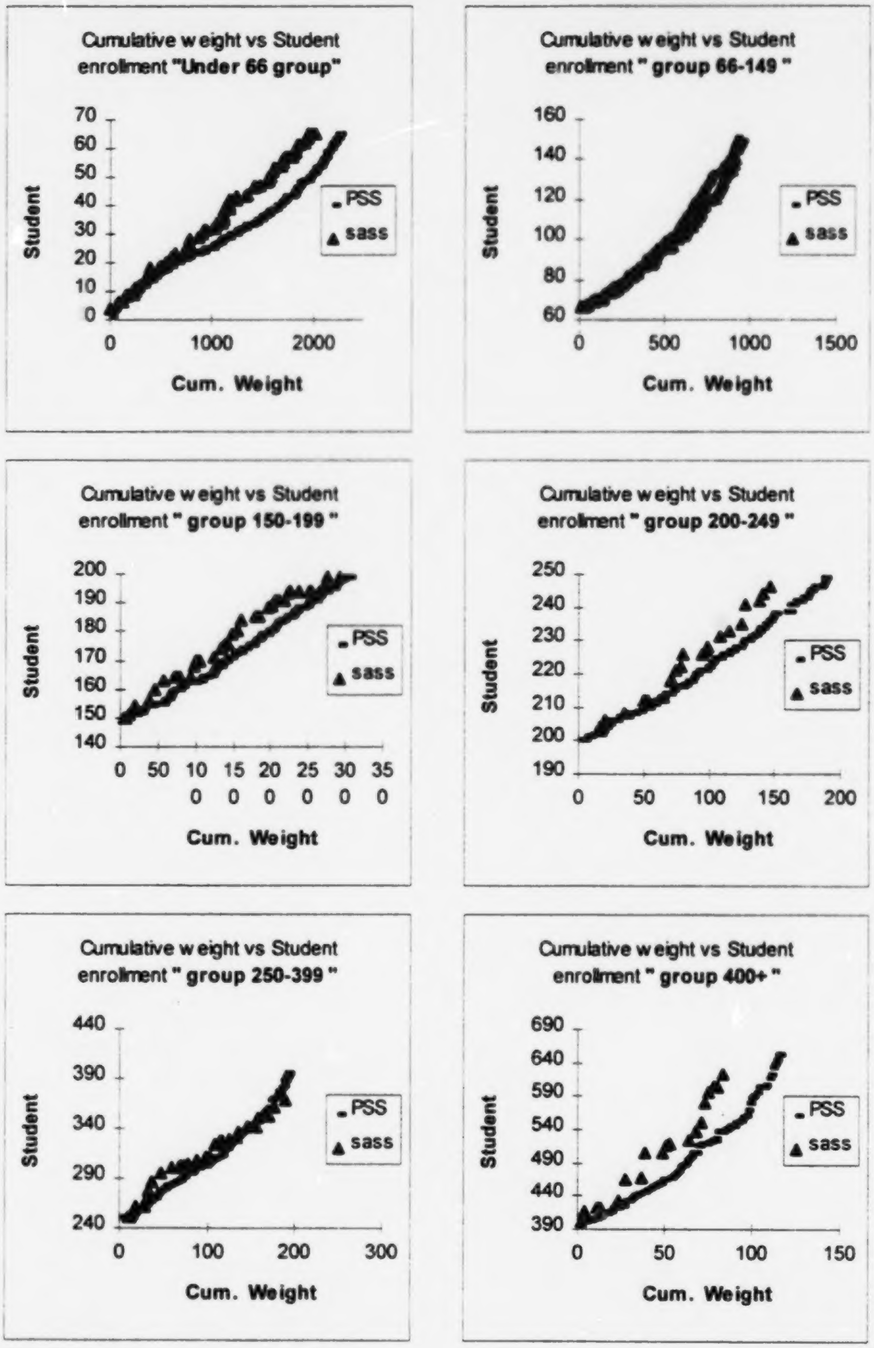
SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 6.5 – Other Unaffiliated: Cumulative weight of the Original SASS weight *before* truncation



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 6.6 – Other Unaffiliated: Cumulative weight of the original SASS weight *after* truncation


Cumulative weight vs Student enrollment "Under 66 group"

Cumulative weight vs Student enrollment " group 66-149 "

Cumulative weight vs Student enrollment " group 150-199 "

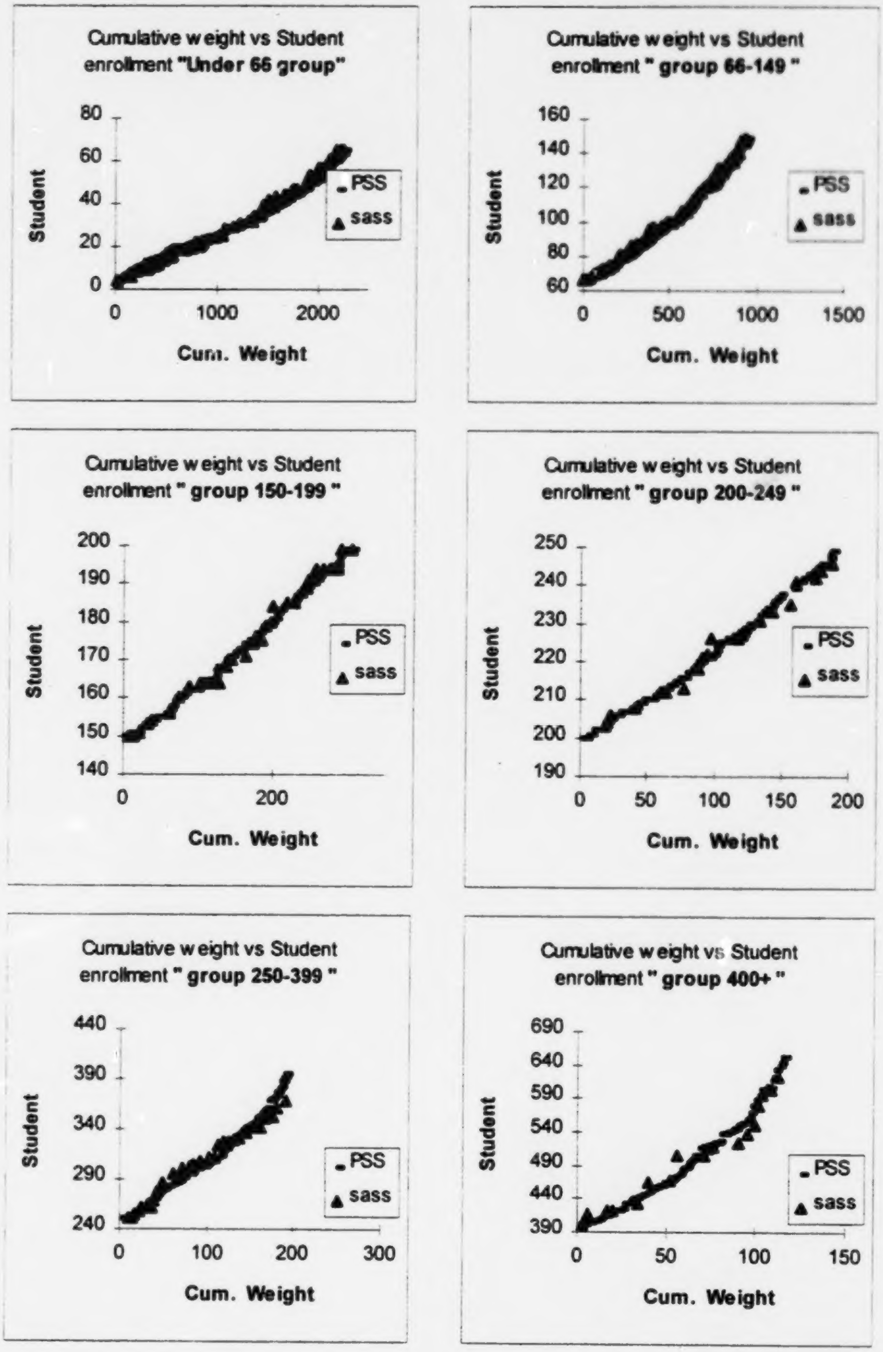
Cumulative weight vs Student enrollment " group 200-249 "

Cumulative weight vs Student enrollment " group 250-399 "

Cumulative weight vs Student enrollment " group 400+ "

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 6.7 -- Other Unaffiliated: Cumulative weight of Olkin GLS after truncating the SASS weights


Cumulative weight vs Student enrollment "Under 66 group"

Cumulative weight vs Student enrollment " group 66-149 "

Cumulative weight vs Student enrollment " group 150-199 "

Cumulative weight vs Student enrollment " group 200-249 "

Cumulative weight vs Student enrollment " group 250-399 "

Cumulative weight vs Student enrollment " group 400+ "

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 6.3 – Other Unaffiliated: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part I - PSS total (3,141 schools)					
1-149	School	831	832	1,535	3,198
	Teacher	5,586	5,023	6,564	17,173
	Student	52,790	48,783	62,774	164,347
150-299	School	235	194	146	575
	Teacher	3,740	2,955	2,323	9,019
	Student	47,574	39,493	30,048	117,114
300-499	School	73	63	37	173
	Teacher	2,141	1,894	1,041	5,076
	Student	27,333	24,373	13,444	65,150
500-749	School	29	12	7	48
	Teacher	1,454	564	238	2,255
	Student	16,574	6,479	3,503	26,556
750+	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	1,168	1,101	1,724	3,994
	Teacher	12,921	10,436	10,166	33,523
	Student	144,271	119,128	109,768	373,168
Part II - Original SASS total (313 schools)					
1-149	School	681	818	1,633	3,133
	Teacher	4,084	4,133	7,599	15,816
	Student	41,306	41,700	85,257	168,263
150-299	School	267	163	103	533
	Teacher	4,035	2,104	1,842	7,981
	Student	59,011	30,274	20,966	110,250
300-499	School	68	87	65	219
	Teacher	1,413	2,070	1,606	5,089
	Student	23,020	30,743	22,760	76,523
500-749	School	33	2	16	52
	Teacher	1,701	168	470	2,338
	Student	18,423	1,232	8,098	27,753
750+	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	1,050	1,071	1,816	3,937
	Teacher	11,234	8,474	11,516	31,225
	Student	141,759	103,949	137,081	382,788

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

BEST COPY AVAILABLE

Table 6.3 -- Other Unaffiliated: Estimates by school size and community type, PSS and SASS compared (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	

Part III - Olkin GLS SASS total (313 schools)

School Size		Central City	Urban Fringe / Large Town	Rural / Small Town	Total
1-149	School	794	792	1,612	3,198
	Teacher	4,992	4,644	7,537	17,173
	Student	45,671	40,459	78,217	164,347
150-299	School	234	199	121	553
	Teacher	3,381	2,921	2,376	8,677
	Student	49,256	36,565	25,088	110,909
300-499	School	43	94	46	183
	Teacher	1,052	2,467	1,535	5,055
	Student	15,538	34,062	16,345	65,945
500-749	School	40	2	18	59
	Teacher	1,939	132	548	2,618
	Student	22,040	967	8,960	31,966
750+	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	1,111	1,086	1,797	3,994
	Teacher	11,363	10,164	11,996	33,523
	Student	132,505	112,053	128,610	373,168

Part IV - Basic GLS SASS total (313 schools)

School Size		Central City	Urban Fringe / Large Town	Rural / Small Town	Total
1-149	School	718	850	1,672	3,239
	Teacher	4,489	4,501	8,000	16,990
	Student	43,462	43,413	86,863	173,738
150-299	School	245	160	112	517
	Teacher	3,923	2,320	2,377	8,621
	Student	54,533	29,410	23,194	107,137
300-499	School	57	70	67	195
	Teacher	1,211	2,018	2,195	5,424
	Student	19,542	25,068	24,175	68,784
500-749	School	24	7	11	43
	Teacher	1,592	559	337	2,488
	Student	13,599	4,103	5,808	23,509
750+	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	1,044	1,087	1,862	3,994
	Teacher	11,217	9,397	12,909	33,523
	Student	131,136	101,993	140,040	373,168

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 6.4 -- Other Unaffiliated: Estimates by school size and community type, PSS and SASS compared in percent difference

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and original SASS totals					
1-149	School	18.00	1.65	-6.41	2.03
	Teacher	26.88	17.72	-15.77	7.90
	Student	21.76	14.52	-35.82	-2.38
150-299	School	-13.69	16.06	29.60	7.33
	Teacher	-7.88	28.82	20.71	11.51
	Student	-24.04	23.34	30.23	5.86
300-499	School	7.20	-37.47	-75.44	-26.77
	Teacher	33.99	-9.31	-54.24	-0.26
	Student	15.78	-26.14	-69.29	-17.46
500-749	School	-13.30	81.66	-145.18	-8.33
	Teacher	-17.03	70.22	-97.31	-3.69
	Student	-11.15	80.98	-131.18	-4.51
750+	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	10.16	2.78	-5.36	1.42
%diff	Teacher	13.06	18.79	-13.28	6.86
From PSS	Student	1.74	12.74	-24.88	-2.58
Percent difference from PSS and Olkin GLS SASS totals					
1-149	School	4.45	4.85	-5.04	0.00
	Teacher	10.64	7.53	-14.82	0.00
	Student	13.49	17.06	-24.60	0.00
150-299	School	0.58	-2.31	17.00	3.77
	Teacher	9.61	1.17	-2.26	3.79
	Student	-3.54	7.41	16.51	5.30
300-499	School	41.03	-47.54	-25.62	-5.63
	Teacher	50.86	-30.28	-47.46	0.42
	Student	43.15	-39.76	-21.58	-1.22
500-749	School	-36.14	85.55	-170.60	-25.02
	Teacher	-33.39	76.65	-130.22	-16.10
	Student	-32.98	85.08	-155.77	-20.37
750+	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	4.92	1.42	-4.24	0.00
%diff	Teacher	12.06	2.60	-17.99	0.00
From PSS	Student	8.16	5.94	-17.16	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 6.4 – Other Unaffiliated: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
		Percent difference from PSS and Basic GLS SASS totals			
1-149	School	13.66	-2.10	-8.93	-1.28
	Teacher	19.64	10.39	-21.88	1.06
	Student	17.67	11.01	-38.37	-5.71
150-299	School	-4.34	17.62	23.37	10.09
	Teacher	-4.90	21.51	-2.33	4.42
	Student	-14.63	25.53	22.81	8.52
300-499	School	21.72	-11.05	-82.42	-12.46
	Teacher	43.43	-6.55	-110.81	-6.85
	Student	28.51	-2.85	-79.82	-5.58
500-749	School	17.17	38.84	-75.80	9.67
	Teacher	-9.56	0.87	-41.60	-10.33
	Student	17.95	36.68	-65.79	11.47
750+	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total %diff From PSS	School	10.62	1.29	-8.02	0.00
	Teacher	13.19	9.95	-26.98	0.00
	Student	9.10	14.38	-27.58	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

3.7 NONSECTARIAN REGULAR TYPOLOGY

The Nonsectarian Regular typology is a fairly small proportion of all private schools. For example, in the 1993-94 Private School Survey, there were an estimated 2,198 Nonsectarian Regular schools or just about 8% of the private school total for that year.

In table 7.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the SASS school total is much larger than the PSS (by about 15.7%); SASS, also, estimates many more teachers and students than are shown in PSS (8.0% and 12.8% more respectively).

To set the stage for the calculations that follow, it might be worth looking at figure 7.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black □'s. Overlaying the PSS data and shown as gray Δ's are the data from the 301 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS and along essentially the same axis. In fact, the slopes of the student/teacher relationship are similar, but clearly distinguishable, being 8.6 for PSS and 8.2 for SASS. While not directly comparable because of differences in sample designs, the R^2 values for the student/teacher ratios in both (unweighted) samples show analytically what can be observed from the graph, namely that the scatter is fairly modest around the average teacher/student relationship. These values are $R^2 = .82$ (PSS) and $R^2 = .81$ (SASS).

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing these new GLS estimators, a decision was first made about which sample cases to use (see section 3.7.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.7.2). The results of the Basic GLS were also obtained (section 3.7.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and Basic GLS versions (section 3.7.4). An independent assessment (section 3.7.5) concludes the discussion.

3.7.1 Determining Outliers. -- Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 7.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

Table 7.1 – Nonsectarian Regular: Weighted schools totals before excluding outliers (Based on 1,856 PSS and 301 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	2,198	2,544	-346
Teachers	51,748	55,911	-4,163
Students	481,423	542,980	-61,557

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

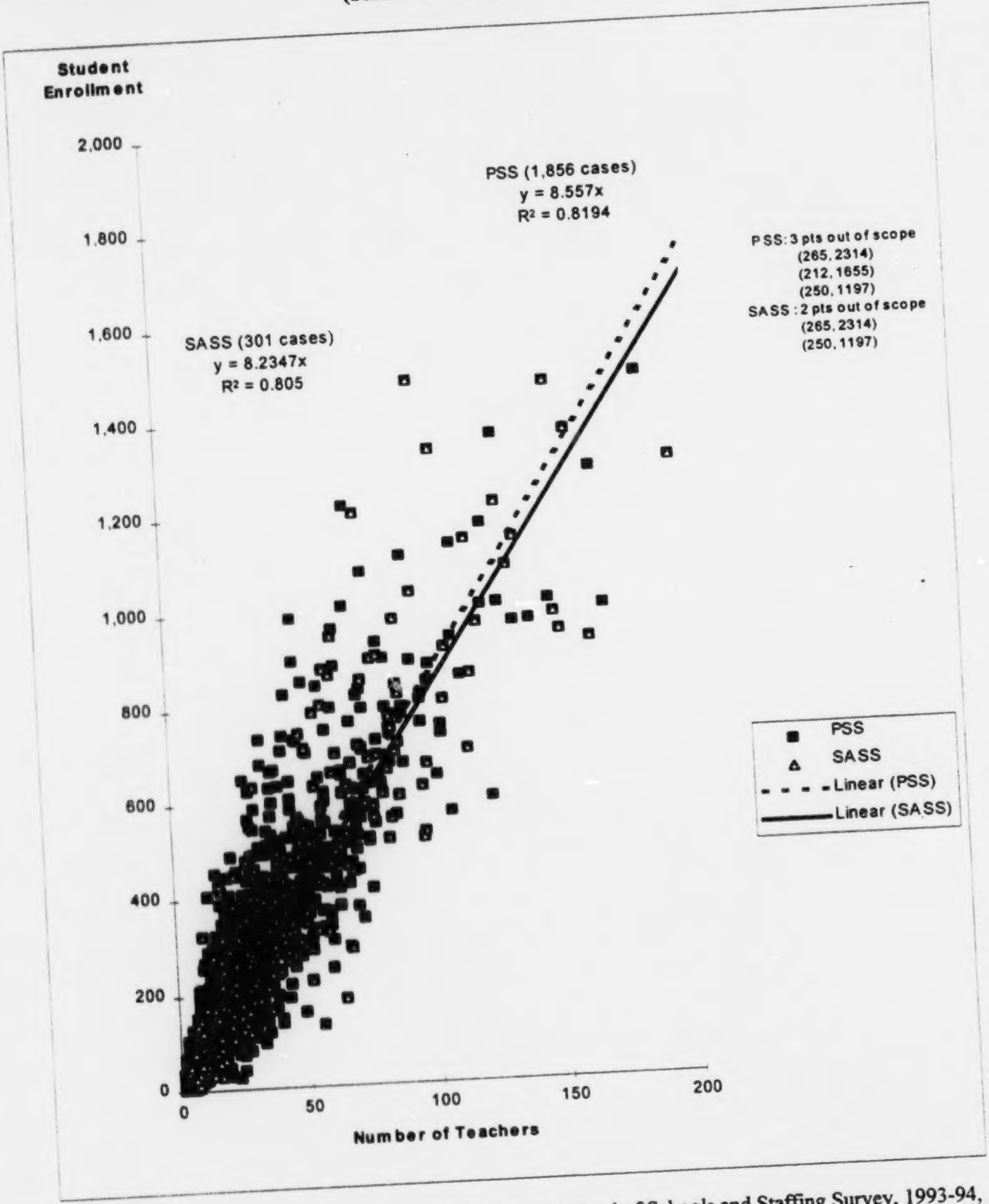
Table 7.2 – Nonsectarian Regular: Weighted schools totals after excluding outliers (Based on 1,839 PSS and 288 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	2,179	2,530	-351
Teachers	49,398	52,904	-3,506
Students	460,151	513,741	-53,590

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 7.1 – Nonsectarian Regular: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

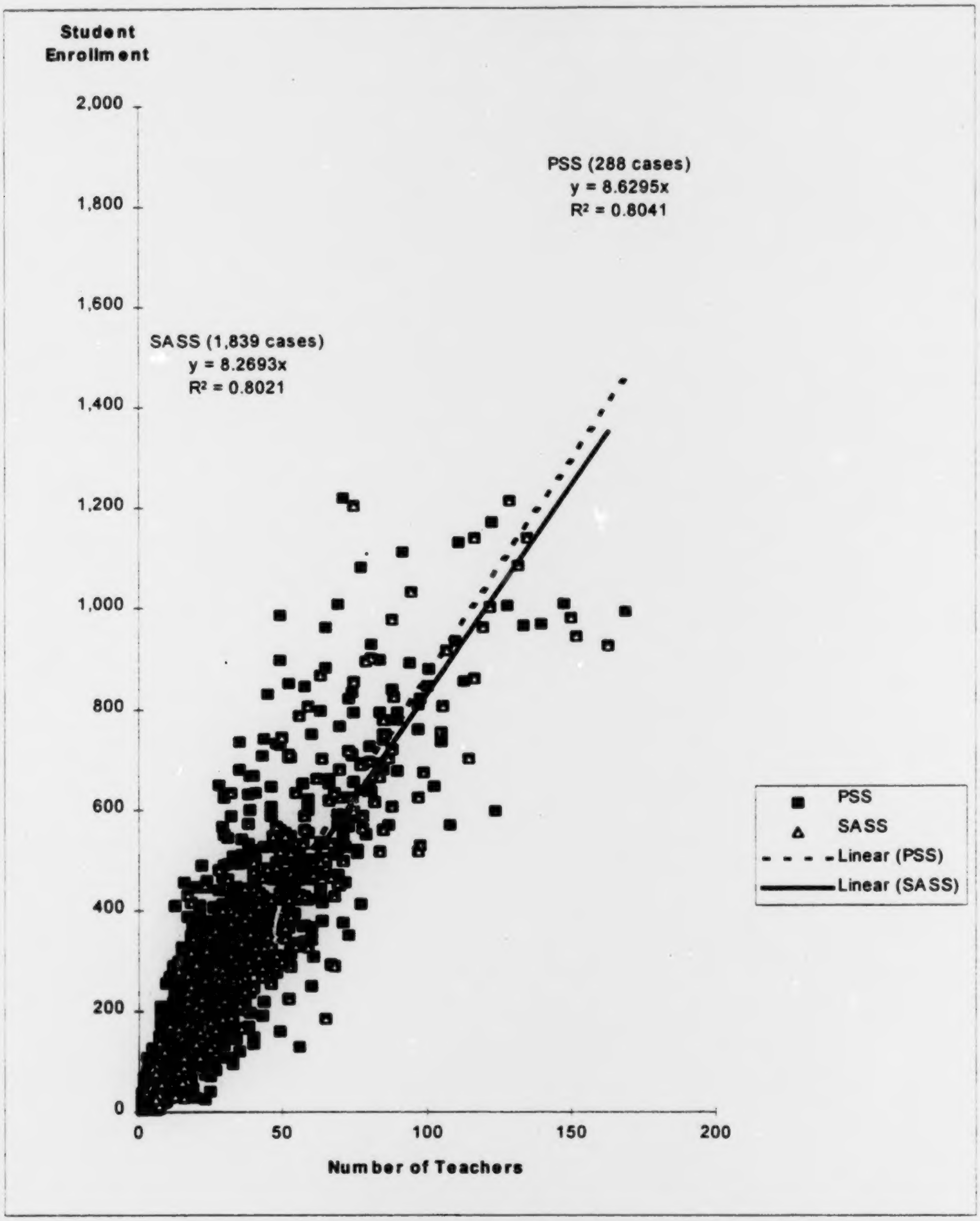
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 7.2 -- Nonsectarian Regular: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

For the Nonsectarian Regular typology, simple visual inspection resulted in reducing the PSS sample by 11 cases -- with a corresponding reduction in the SASS sample of 7 cases. The visual inspection, though, had to be supplemented by a more analytic method, which systematically excluded points more than a certain distance from the overall center of the combined PSS/SASS samples. After this second step, there was a further reduction of 6 PSS and 6 SASS points.

Figure 7.2 is the plot of the remaining 1839 PSS and 288 SASS cases. Notice that the student/teacher relationships have not changed appreciably from those in figure 7.1. In particular, the student/teacher ratio in PSS went from 8.557 to 8.629; for SASS the ratio went from 8.235 to 8.269. This seemed close enough for the GLS method to have a chance of working without negative weights. Notice further, the scatter of points in both samples is visually perhaps a little tighter, albeit the R^2 values changed almost not at all.

3.7.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\lambda} = \mathbf{M}^{-1}\underline{d}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. The starting point was the new typology totals for PSS and SASS. These are shown in table 7.2 below.

To carry out the Olkin GLS, the schools were placed into four school size classes (specifically under 150, 150 to 499, 500 to 749, 750 and above).

When initially run, though, the Olkin GLS resulted in 26 cases with excessively small weights (between .2 and .7). In these instances, a Winsorizing constant of .5 was added and then the Olkin adjustment was redone. Anyway, after both these steps, the difference between the PSS and SASS estimates had shrunk (considerably in overall absolute value) to

$$\underline{d} = \begin{array}{r} -86 \\ 1137 \\ -8798 \end{array}$$

The matrix **M** was obtained by tabulating the 1993-94 SASS file for the 288 schools remaining in the SASS sample. The values are

$$\begin{array}{rrr} 288 & 11403 & 98970 \\ 11403 & 735175 & 6079382 \\ 98970 & 6079382 & 54284736 \end{array}$$

Solving for $\underline{\lambda}$ yields

$$\underline{\lambda}' = (-0.77597, +0.04361, -0.00363)$$

and the Olkin GLS weights are of the form

$$u_i = w_i - 0.77597 + 0.04361t_i - 0.00363s_i$$

Notice that all the original weights are lowered considerably (by about .8); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered further (usually they would not be lowered, except for the schools with the very largest enrollments). These additional school-by-school adjustments appear to be small and should not have much if any adverse consequences. One final comment, while the values for $\underline{\lambda}$ are only shown to six significant digits, the calculations have been carried out in double precision.

3.7.3 Basic GLS Procedure Employed. -- To carry out the basic GLS weighting the equation

$$\underline{\lambda} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 7.2 above that $\underline{\mathbf{d}}$ for the basic GLS would be

$$\underline{\mathbf{d}} = \begin{matrix} -351 \\ -3506 \\ -53590 \end{matrix}$$

The matrix **M** is again obtained by tabulating the eligible portion of the 1993-94 SASS file for the Nonsectarian Regular schools in the SASS sample. Because the Olkin and Basic GLS employ just the unweighted sample to calculate **M**, it is the same for both (and hence not shown).

Solving for $\underline{\lambda}$ yields this time

$$\underline{\lambda}' = (-2.44982, +0.05968, -0.00316)$$

and the basic GLS weights are of the form

$$u_i = w_i - 2.44982 + 0.05968t_i - 0.00316s_i$$

Notice that again the original weights are lowered, this time by quite a bit more than the amount that the Olkin GLS weights were; again, however,, depending on the teacher and student counts in the sampled school, they may be increased again or lowered further. Looking just at the equation, concerns about negative weights arise and, indeed, these did materialize.

3.7.4 Operational Characteristics. -- To examine the Basic and Olkin GLS reweighting done, several "diagnostics" will be looked at. One statistic that may merit immediate attention is what happened to the spread in the weights themselves. Did the spread grow larger or smaller?

Figure 7.3 provides this information in its upper panels, which compare the original and two GLS adjustments. The original (or y) weights in the upper panel have a slightly smaller spread than the GLS (or x) weights (since the equation which fits them is $y = 1.0442x$). Now, on the other hand, the original SASS weights are related to the Olkin GLS weights by an equation of the form $y = 1.1593x$. Hence, the spread in the Olkin (or x) weights is considerably smaller than for the original (or y) weights.

The R^2 values shown in the upper panel in figures 7.3 might be commented on too. Both are quite high, at or above 0.99 and most of the points lie very close to or just above the 45 degree line. The problem of negative weights did not arise for the Olkin GLS method. There were, though, 41 schools with weights close too but still less than one; for the Basic GLS, there were 42 negative weights and 16 more less than one.

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 7.3 will continue to be our source. This time, though, look at the panel in the second row of the scatterplot matrix where the Basic and Olkin GLS are graphed. The regression average of the basic GLS (or y) values is larger than the corresponding average of the Olkin GLS (or x) values. The R^2 of .99+ between the two methods suggests a closeness between them, despite the problem of negative weights for the Basic GLS.

3.7.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Nonsectarian Regular Typology are available in tables 7.3 and 7.4, plus figure 7.4:

-- Table 7.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.

-- Table 7.4 is based on table 7.3 but focuses directly on percentage differences between the three SASS estimates and PSS.

-- Figure 7.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Note that this graph is in logs.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 7.3 and especially 7.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does reasonably well. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS half of the time. It is closer than the Basic GLS in 11 out of 18 comparisons). The results by community type are good for the Olkin GLS relative to the other two. This is unexpected, since the Olkin approach did not try to control by community type (as it had by school size).

In figure 7.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already very close to the PSS, since the best fit regression equation which connects the various estimates is $y = .9908x$. There is just a little roughness around this average, however, as displayed visually and summarized by the R^2 value which is equal to $R^2 = .9912$. This is extremely good, suggesting that the SASS sample of Nonsectarian Regular schools is excellent.

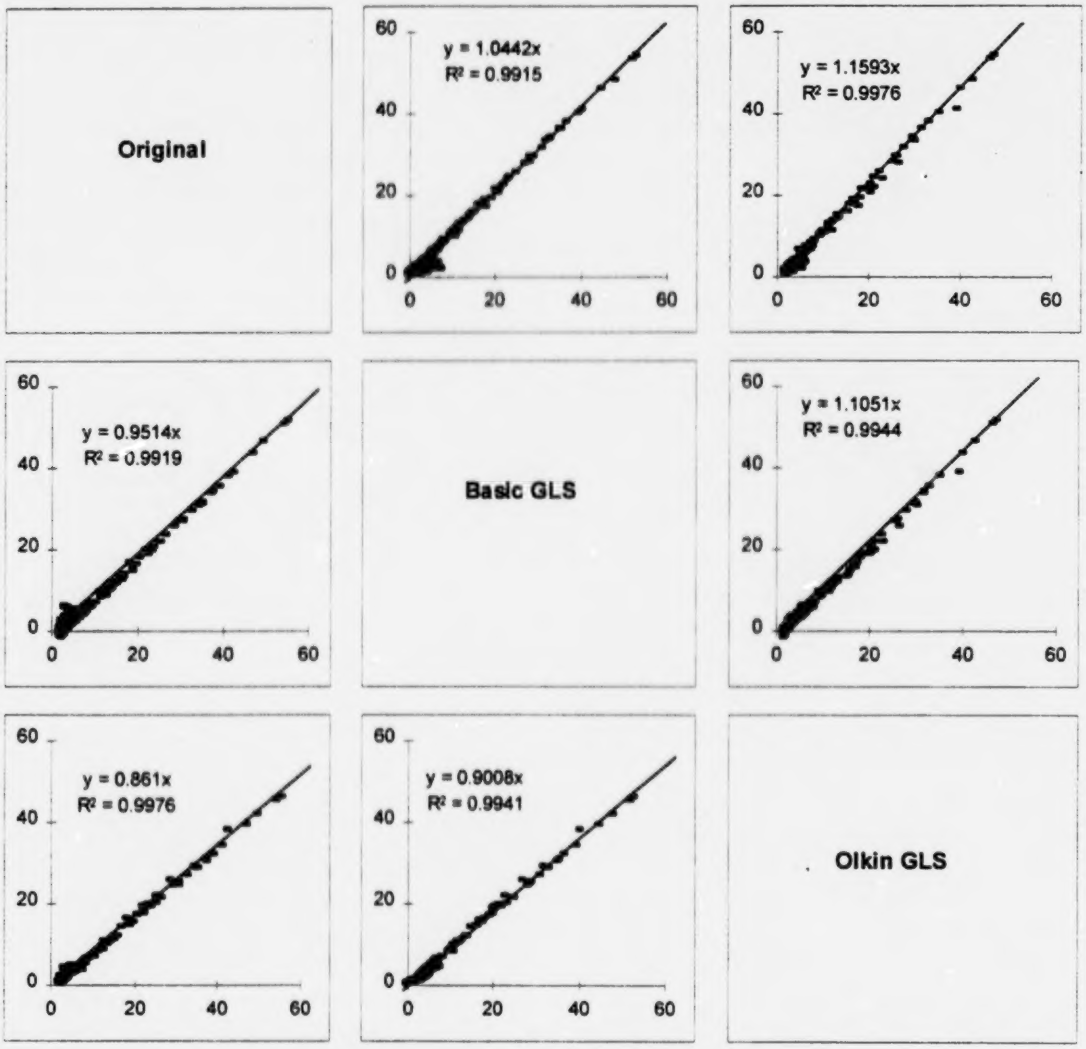
As in figure 7.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit is similar, yielding the relationship $y = 1.0041x$. Again, the average results for this method remain good. A great deal more roughness is exhibited around the average, as evidenced by the much lower R^2 value in this case ($R^2 = .9804$).

Finally, in figure 7.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields the relationship $y = 1.0037x$. The average results for this method are intermediate between the other two, with again an excellent R^2 value in this case of $R^2 = .9883$.

What can be concluded about this typology? The Olkin GLS method seems in no way inferior overall to the original SASS weighted file. To its credit, moreover, it hits the overall PSS school, teacher, and student totals exactly. The Olkin GLS method, furthermore and not surprisingly, does as well or better than the other two when estimates are looked at by school size. The Basic GLS method, while good in some respects, cannot be used without further adjustment because of the negative weights which exist.

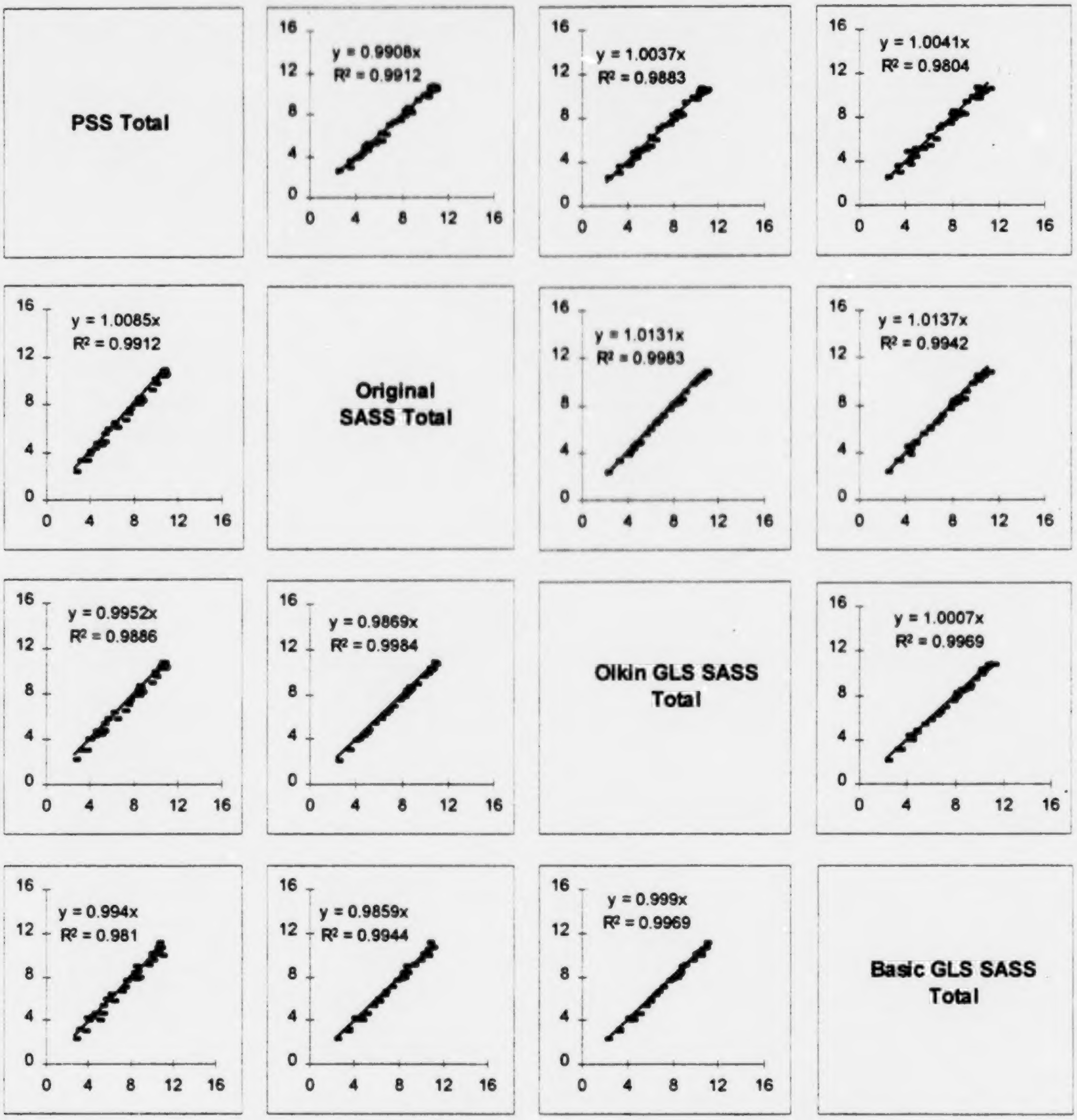
In the summary and recommendations section, additional comments will be made about how the Olkin GLS might be improved further, leading to still better results. (Also see Kaufman, Li, and Scheuren 1995.)

Figure 7.3 -- Nonsectarian Regular: Scatterplot matrix comparing original, basic GLS, and Olkin GLS weights



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 7.4 – Nonsectarian Regular: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and basic GLS SASS totals by school size and community type from Table 7.3 (in log scale)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 7.3 – Nonsectarian Regular: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part I - PSS total (1,839 schools)					
1 - 149	School	502	403	243	1,147
	Teacher	3,811	3,362	2,410	9,583
	Student	31,266	27,090	17,031	75,386
150 - 299	School	173	156	168	497
	Teacher	3,698	3,821	3,988	11,506
	Student	36,548	33,353	37,640	107,540
300 - 499	School	124	88	98	310
	Teacher	5,119	3,513	3,394	12,026
	Student	47,288	33,974	37,382	118,644
500 - 749	School	68	50	33	151
	Teacher	4,571	3,337	1,477	9,385
	Student	41,468	30,262	19,596	91,327
750 +	School	41	19	14	74
	Teacher	4,217	1,650	1,032	6,899
	Student	38,312	16,470	12,472	67,254
Total	School	908	716	555	2,179
	Teacher	21,415	15,683	12,300	49,398
	Student	194,882	141,150	124,120	460,151
Part II - Original SASS total (288 schools)					
1 - 149	School	416	640	384	1,439
	Teacher	3,132	3,667	3,174	9,974
	Student	31,889	31,120	23,280	86,289
150 - 299	School	244	137	123	504
	Teacher	5,289	3,488	3,642	12,419
	Student	52,325	30,184	30,171	112,680
300 - 499	School	94	118	118	330
	Teacher	4,013	4,535	3,689	12,237
	Student	36,481	49,380	49,700	135,561
500 - 749	School	77	57	27	161
	Teacher	5,113	3,872	1,295	10,281
	Student	47,128	35,270	16,983	99,382
750 +	School	47	26	10	83
	Teacher	4,910	2,037	817	7,764
	Student	43,872	24,346	9,256	77,474
Total	School	878	977	662	2,517
	Teacher	22,457	17,600	12,618	52,675
	Student	211,696	170,300	129,390	511,386

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 7.3 – Nonsectarian Regular: Estimates by school size and community type, PSS and SASS compared (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part III - Olkin GLS SASS total (288 schools)					
1 - 149	School	342	537	325	1,205
	Teacher	2,569	3,051	2,688	8,308
	Student	26,340	25,765	19,592	71,697
150 - 299	School	221	123	114	458
	Teacher	4,892	3,229	3,502	11,623
	Student	47,344	27,324	28,117	102,786
300 - 499	School	78	105	100	283
	Teacher	3,435	4,180	3,175	10,790
	Student	30,337	43,953	42,282	116,572
500 - 749	School	73	56	22	152
	Teacher	5,107	4,095	1,091	10,293
	Student	44,597	35,068	13,947	93,611
750 +	School	51	21	9	81
	Teacher	5,888	1,822	673	8,384
	Student	48,373	19,486	7,624	75,484
Total	School	765	843	570	2,179
	Teacher	21,892	16,377	11,129	49,398
	Student	196,992	151,596	111,563	460,151
Part IV - Basic GLS SASS total (288 schools)					
1 - 149	School	342	579	349	1,270
	Teacher	2,481	3,149	2,799	8,428
	Student	26,910	26,830	20,545	74,284
150 - 299	School	199	105	98	403
	Teacher	4,329	2,747	3,036	10,113
	Student	42,571	23,449	24,498	90,518
300 - 499	School	56	99	96	251
	Teacher	2,648	3,909	3,012	9,569
	Student	22,267	41,804	40,953	105,024
500 - 749	School	73	58	22	153
	Teacher	5,168	4,264	1,103	10,534
	Student	44,694	36,111	14,052	94,857
750 +	School	66	26	10	101
	Teacher	7,748	2,233	772	10,753
	Student	63,230	23,488	8,751	95,469
Total	School	737	866	576	2,179
	Teacher	22,374	16,302	10,722	49,398
	Student	199,671	151,681	108,799	460,152

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 7.4 – Nonsectarian Regular: Estimates by school size and community type, PSS and SASS compared in percent difference

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and original SASS totals					
1 - 149	School	17.07	-58.91	-57.88	-25.47
	Teacher	17.82	-9.09	-31.73	-4.08
	Student	-1.99	-14.88	-36.69	-14.46
150 - 299	School	-40.79	12.32	26.60	-1.39
	Teacher	-43.04	8.72	8.66	-7.93
	Student	-43.17	9.50	19.84	-4.78
300 - 499	School	24.00	-33.23	-20.61	-6.39
	Teacher	21.60	-29.11	-8.69	-1.76
	Student	22.85	-45.34	-32.95	-14.26
500 - 749	School	-13.38	-13.26	19.49	-6.13
	Teacher	-11.87	-16.05	12.33	-9.55
	Student	-13.65	-16.55	13.33	-8.82
750 +	School	-13.36	-40.27	24.30	-13.21
	Teacher	-16.45	-23.45	20.81	-12.55
	Student	-14.51	-47.82	25.78	-15.20
Total % diff from PSS	School	3.32	-36.54	-19.17	-15.51
	Teacher	-4.87	-12.23	-2.58	-6.63
	Student	-8.63	-20.65	-4.25	-11.13
Percent difference from PSS and Olkin GLS SASS totals					
1 - 149	School	31.74	-33.46	-33.95	-5.06
	Teacher	32.59	9.25	-11.56	13.30
	Student	15.75	4.89	-15.04	4.89
150 - 299	School	-27.48	20.99	31.80	7.73
	Teacher	-32.30	15.49	12.19	-1.01
	Student	-29.54	18.07	25.30	4.42
300 - 499	School	37.26	-18.76	-2.28	8.81
	Teacher	32.88	-19.00	6.46	10.27
	Student	35.85	-29.37	-13.11	1.75
500 - 749	School	-8.00	-12.68	33.95	-0.34
	Teacher	-11.73	-22.72	26.13	-9.68
	Student	-7.54	-15.88	28.83	-2.50
750 +	School	-23.62	-13.97	37.63	-9.80
	Teacher	-39.64	-10.42	34.77	-21.52
	Student	-26.26	-18.31	38.86	-12.24
Total % diff from PSS	School	15.70	-17.82	-2.71	0.00
	Teacher	-2.23	-4.43	9.52	0.00
	Student	-1.08	-7.40	10.12	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 7.4 – Nonsectarian Regular: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and Basic GLS SASS totals					
1 - 149	School	31.75	-43.71	-43.85	-10.74
	Teacher	34.91	6.33	-16.13	12.05
	Student	13.93	0.96	-20.63	1.46
150 - 299	School	-15.00	32.47	41.30	18.89
	Teacher	-17.08	28.11	23.86	12.11
	Student	-16.48	29.69	34.92	15.83
300 - 499	School	54.89	-11.99	1.73	19.05
	Teacher	48.26	-11.29	11.27	20.42
	Student	52.91	-23.04	-9.56	11.48
500 - 749	School	-7.98	-15.51	33.67	-1.33
	Teacher	-13.06	-27.77	25.32	-12.25
	Student	-7.78	-19.33	28.29	-3.86
750 +	School	-59.89	-36.48	28.40	-37.54
	Teacher	-83.74	-35.33	25.13	-55.88
	Student	-65.04	-42.61	29.83	-41.95
Total % diff from PSS	School	18.85	-21.04	-3.71	0.00
	Teacher	-4.48	-3.95	12.83	0.00
	Student	-2.46	-7.46	12.34	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

3.8 NONSECTARIAN SPECIAL EMPHASIS TYPOLOGY

The Nonsectarian Special Emphasis typology ranks 7th in size among private schools. For example, in the 1993-94 Private School Survey, there were an estimated 2,106 schools or about 8% of the private school total for that year. (Only the Special Education and Private Catholic typologies had smaller numbers of schools).

In table 8.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the SASS school totals quite a bit lower (at 14.3%); SASS, also, has many fewer teachers and students than are shown in PSS (16.7% to 12.9% less).

To set the stage for the calculations that follow, it might be worth looking at figure 8.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black □'s. Overlaying the PSS data and shown as gray Δ's are the data from the 149 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS and along the same axis. In fact, the slope of the student/teacher relationship is 7.5 for both PSS and SASS. While not directly comparable because of differences in sample designs, the R^2 values for the student/teacher ratios in both (unweighted) samples also show a strong relationship -- at $R^2 = .71$ (PSS) and $R^2 = .85$ (SASS).

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing this new estimator, a decision was first made about which sample cases to use (see section 3.8.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.8.2). The results of the basic GLS were also obtained (section 3.8.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and basic GLS versions (section 3.8.4). An independent assessment (section 3.8.5) concludes the discussion.

3.8.1 Determining Outliers. --Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 8.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

For the Special Emphasis typology, simple visual inspection and a later systematic analysis were needed. There were 24 PSS cases set aside for imputation -- 22 visually and two more after analysis; for the SASS, the corresponding values were eight visually and two more analytically. The remaining PSS sample of 1594 was then employed in all the rest of the

Table 8.1 -- Nonsectarian Special Emphasis: Weighted schools totals before excluding outliers (Based on 1,618 PSS and 149 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	2,106	1,805	301
Teachers	20,794	17,321	3,473
Students	163,251	142,180	21,071

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

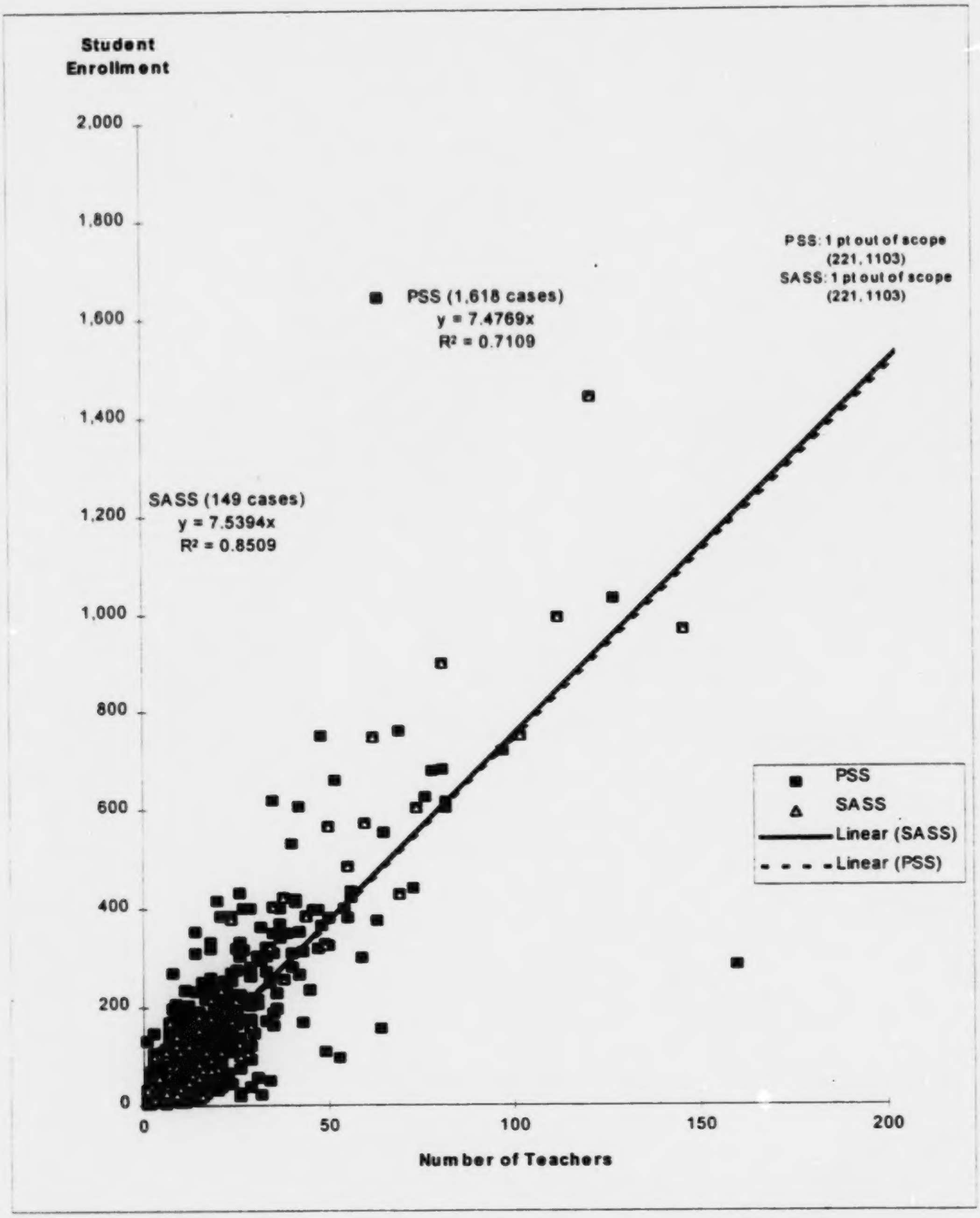
Table 8.2 -- Nonsectarian Special Emphasis: Weighted schools totals after excluding outliers (Based on 1,594 PSS and 139 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	2,079	1,783	296
Teachers	18,431	15,116	3,315
Students	142,627	123,423	19,203

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 8.1 -- Nonsectarian Special Emphasis: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

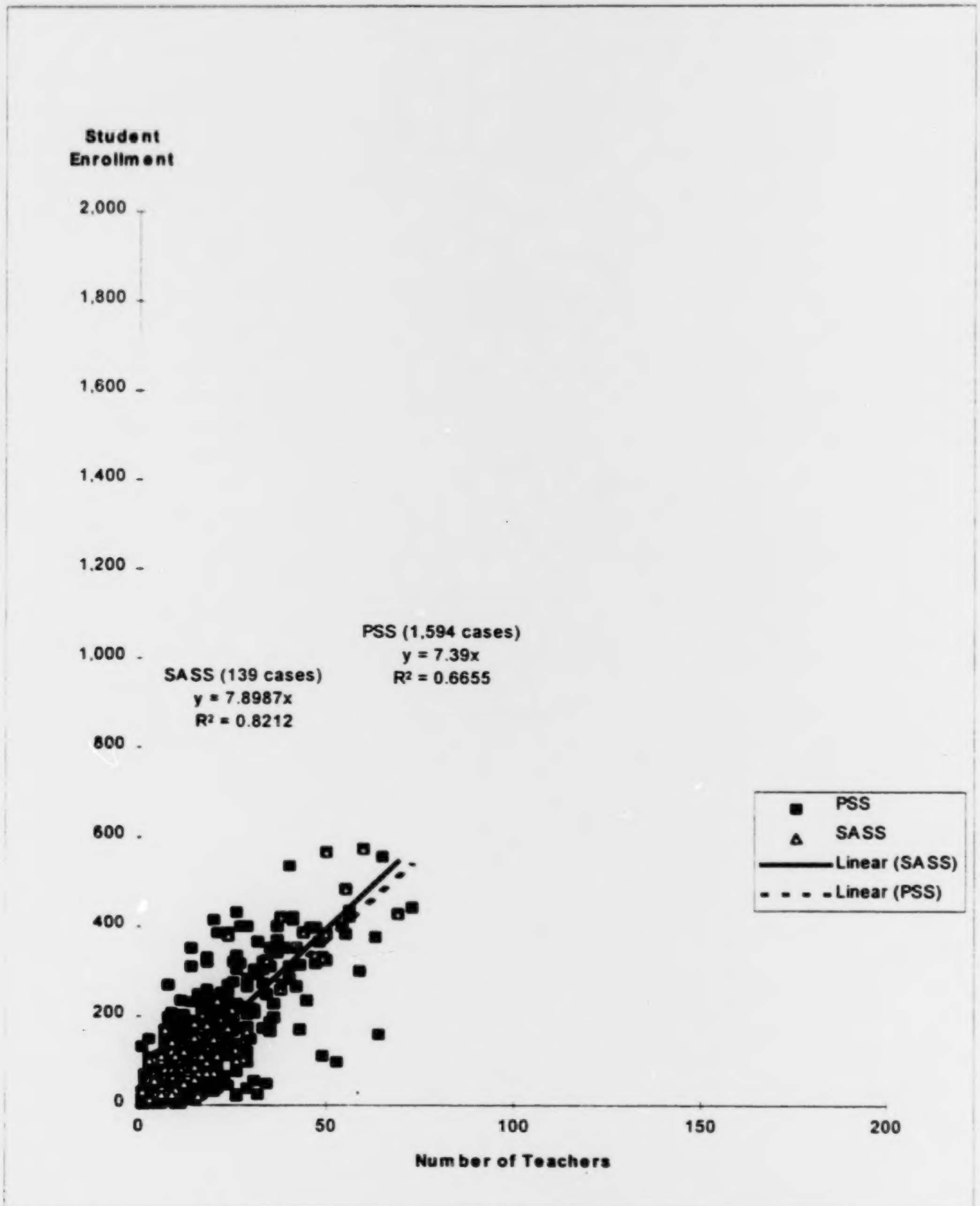
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey:1993-94, Private School Surveys, 1993-94.

Figure 8.2 – Nonsectarian Special Emphasis: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

work commented on here; the corresponding SASS remaining sample was 139. Figure 8.2 is the plot of these remaining cases. Notice that the student/teacher relationships are little changed overall from those in figure 1.1; however, the scatter in both samples appears considerably tighter visually, although the R^2 values are somewhat lower.

3.8.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. The starting points were the new typology totals for PSS and SASS. These are shown in table 8.2 below.

To carry out the Olkin GLS, the schools were grouped into just two school size classes (under 150, 150 and above).

After the Olkin adjustment to each of the two school size groups, the difference between PSS and SASS had shrunk (considerably in overall value) to

$$\underline{\mathbf{d}} = \begin{matrix} -75 \\ 668 \\ -1558 \end{matrix}$$

The matrix **M** was obtained by tabulating the 1993-94 SASS file for the Special Emphasis schools in the SASS sample. The values are

$$\begin{matrix} 139 & 1692 & 13456 \\ 1692 & 41808 & 330228 \\ 13456 & 330228 & 2892692 \end{matrix}$$

Solving for $\underline{\boldsymbol{\lambda}}$ yields

$$\underline{\boldsymbol{\lambda}}' = (-1.41860, +0.25968, -0.02358)$$

and the Olkin GLS weights are of the form

$$u_i = w_i - 1.41860 + 0.25968t_i - 0.02358s_i$$

Notice that all the original weights are lowered by a considerable amount (about 1.4); and, then, depending on the teacher and student counts in the sampled school, they may be lowered further (usually this would not occur except for the very largest schools). These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts are so small. One final comment, while the values for $\underline{\lambda}$ are only shown to six significant digits, the calculations have been carried out in double precision.

3.8.3 Basic GLS Procedure Employed. -- To carry out the basic GLS weighting the equation

$$\underline{\lambda} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 8.2 above that $\underline{\mathbf{d}}$ for the basic GLS would be

$$\underline{\mathbf{d}} = \begin{matrix} 296 \\ 3315 \\ 19203 \end{matrix}$$

The matrix **M** is again obtained by tabulating the eligible portion of the 1993-94 SASS file for the Special Emphasis schools in the SASS sample. Because the Olkin and Basic GLS employ just the unweighted sample to calculate **M**, it is the same for both (and hence not shown).

Solving for $\underline{\lambda}$ yields this time

$$\underline{\lambda}' = (+2.33752, +0.18476, -0.02532)$$

and the basic GLS weights are of the form

$$u_i = w_i + 2.33752 + 0.18476t_i - 0.02532s_i$$

Notice that all the original weights this time are raised (and by a considerable amount); then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered (usually they would not be "lowered" except for the very largest schools). These additional school-by-school adjustments appear important because of the size of the coefficients employed. Looking just at the equation, concerns about negative weights might arise but, as will be seen below, these did not materialize.

3.8.4 Operational Characteristics. -- Both the Basic and Olkin GLS reweighting done, as described above, seems to have worked well. To indicate why this observation is made, several "diagnostics" will be looked at. One statistic that may merit immediate attention is what happened to the spread in the weights themselves. Did the spread grow larger or smaller?

Figure 8.3 provides this information in its upper panels, which compare the original and two GLS adjustments. Both GLS weights have a wider spread than does the original SASS weight. This follows by noting that the original (or y) weights in the upper panel are related to the GLS (or x) weights by the expressions y = .9252x (Basic GLS); and y = .8384x (Olkin GLS).

The R^2 values shown in the upper panel in figures 8.3 might be commented on too, along with the appearance of the scatter itself. In particular, very little evidence exists to indicate that either of the reweighting approaches altered the original weights in any major way, beyond rescaling them. The R^2 values are both above 0.98 and most of the points lie on or just about the 45 degree line. The problem of negative weights did not arise either and there was only a few cases where the GLS weights were less than one (four for the Olkin GLS and one for the Basic GLS).

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 8.3 will continue to be our source. This time, though, look at panel in the second row of the scatterplot matrix where the Basic and Olkin GLS are graphed. The basic GLS (or y) values are somewhat smaller than the Olkin GLS (or x) values. Beyond this rescaling, there is virtually no difference in the weights -- as evidenced by an R^2 of .98 between the two methods. The plotted points confirm this.

3.8.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Catholic Parochial Typology are available in tables 8.3 and 8.4, plus figure 8.4:

-- Table 8.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.

-- Table 8.4 is based on table 8.3 but focuses directly on percentage differences between the three SASS estimates and PSS.

-- Figure 8.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Notice the graph has been plotted in logs.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 8.3 and especially 8.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does the best on the average. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS 10 times and closer than the Basic GLS in eight out of 18 times. The data by community type also appear better for the Olkin GLS, even though the Olkin approach did not try to control by community type, as it had by school size.

In figure 8.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already very close to the PSS, since the best fit regression equation that connects the various estimates is y = 1.0183x. There is some roughness around this average, however, as displayed visually and summarized by the R^2 value that is equal to $R^2 = .9577$. This is quite good, suggesting that the SASS sample of Special Emphasis schools is excellent.

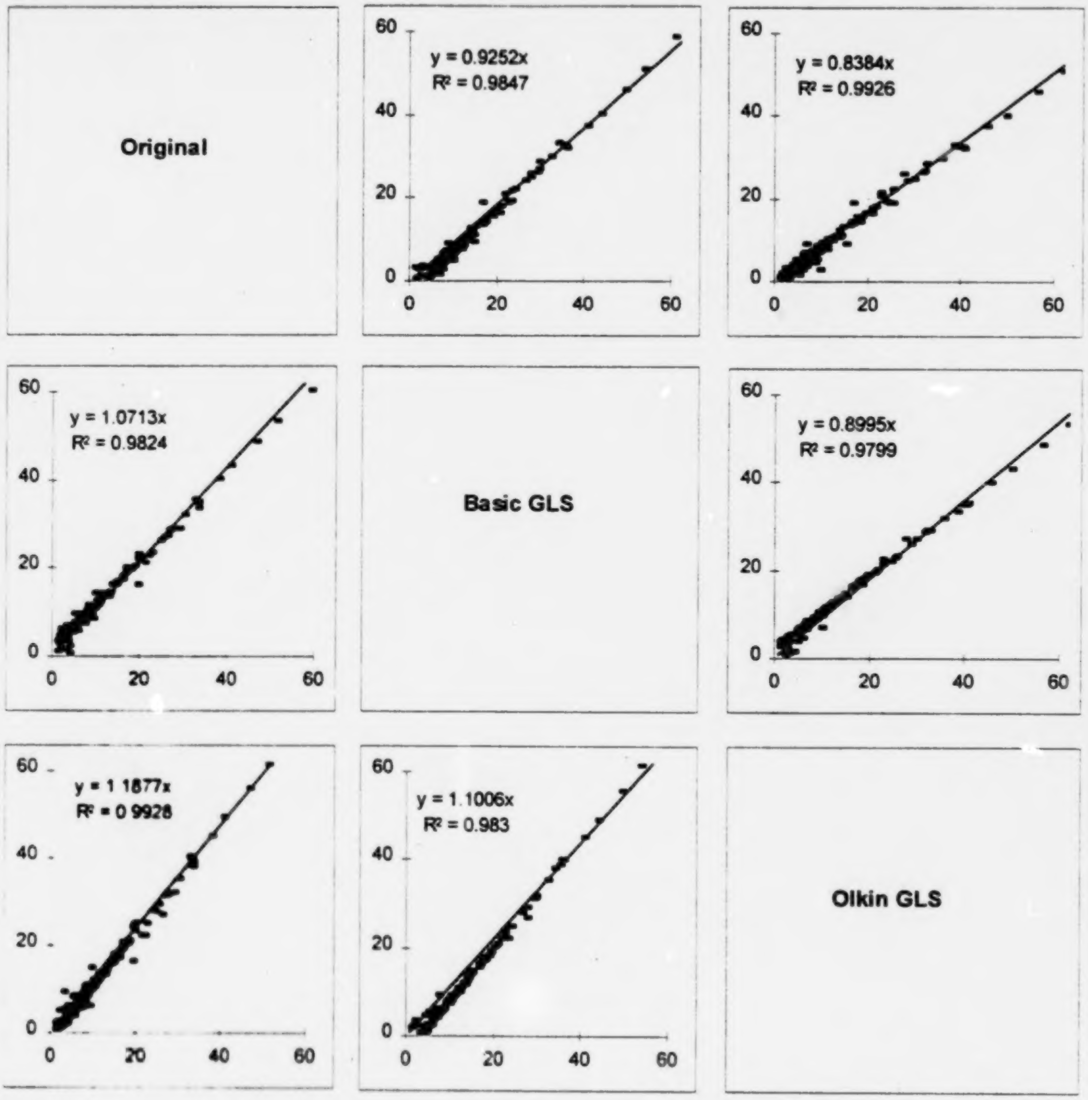
As in figure 8.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit yields the relationship y = 1.0119x. The average results for this method remain good. Again, there is little roughness exhibited around the regression average, as evidenced by the slightly larger R^2 value in this case ($R^2 = .9502$).

Finally, in figure 8.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields the relationship y = 1.0026x. The average results for this method are slightly better than the other two. The roughness exhibited around the regression average is similar to the other two estimators, as evidenced by the R^2 value of $R^2 = .9539$.

What can be concluded about this typology? Neither the Basic nor the Olkin GLS methods seem in any way inferior overall to the original SASS weighted file. To their credit, both hit the PSS school, teacher, and student totals exactly. The Olkin GLS method, furthermore and not surprisingly, does as well or better than the other two when estimates are looked at by school size (and even community type).

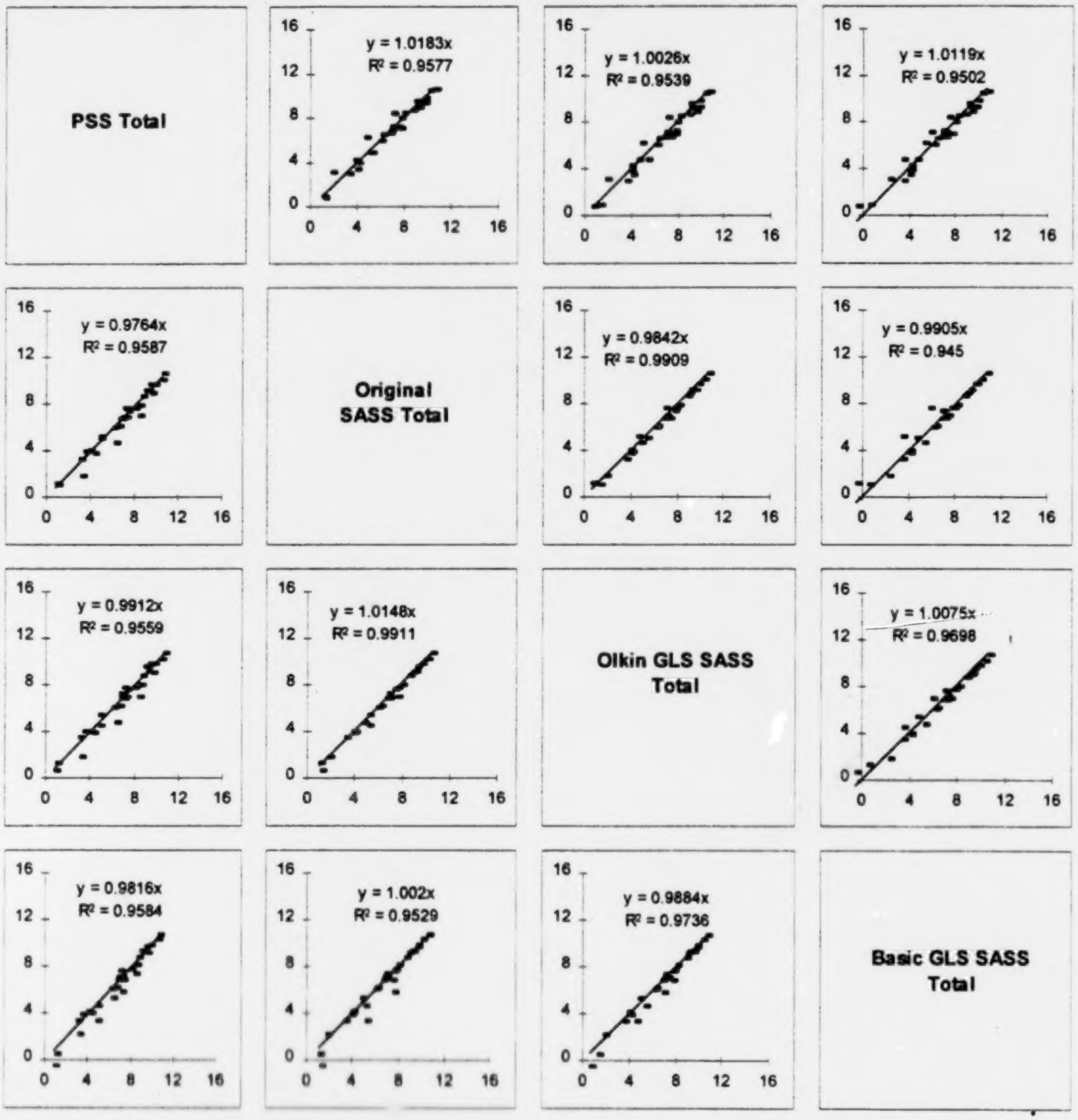
In the summary and recommendations section, comments will be made about how the Olkin GLS might be improved further, leading to still better results.

Figure 8.3 – Nonsectarian Special Emphasis: Scatterplot matrix comparing original, basic GLS, and Olkin GLS weights



SOURCE U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 8.4 – Nonsectarian Special Emphasis: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and basic GLS SASS totals by school size and community type from Table 8.3 (in log scale)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 8.3 -- Nonsectarian Special Emphasis: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part I - PSS total (1,594 schools)					
1 - 149	School	771	694	412	1,877
	Teacher	5,614	4,702	2,898	13,214
	Student	40,525	33,701	17,418	91,645
150 - 299	School	68	50	22	141
	Teacher	1,262	1,047	509	2,819
	Student	13,499	10,059	4,394	27,952
300 - 499	School	32	20	5	56
	Teacher	1,085	867	197	2,149
	Student	11,128	7,371	1,901	20,400
500 - 749	School	2	3	NA	5
	Teacher	125	123	NA	249
	Student	1,227	1,403	NA	2,630
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	873	767	439	2,079
	Teacher	8,087	6,740	3,604	18,431
	Student	66,379	52,535	23,713	142,627
Part II - Original SASS total (139 schools)					
1 - 149	School	779	439	382	1,600
	Teacher	5,413	2,661	2,022	10,096
	Student	36,951	24,280	15,361	76,592
150 - 299	School	44	53	6	103
	Teacher	912	904	110	1,926
	Student	7,296	9,432	1,030	17,757
300 - 499	School	47	26	NA	72
	Teacher	1,922	840	NA	2,762
	Student	15,708	9,910	NA	25,619
500 - 749	School	3	3	NA	6
	Teacher	168	163	NA	330
	Student	1,900	1,549	NA	3,449
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	873	521	388	1,782
	Teacher	8,414	4,567	2,133	15,114
	Student	61,855	45,171	16,390	123,417

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 8.3 -- Nonsectarian Special Emphasis: Estimates by school size and community type, PSS and SASS compared (cont'd)

Part III - Olkin GLS SASS total (139 schools)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
1 - 149	School	919	508	451	1,879
	Teacher	6,675	3,164	2,443	12,281
	Student	43,835	27,661	18,099	89,594
150 - 299	School	50	52	6	108
	Teacher	1,063	979	118	2,160
	Student	8,117	9,441	1,057	18,616
300 - 499	School	53	34	NA	87
	Teacher	2,225	1,457	NA	3,682
	Student	17,766	13,556	NA	31,322
500 - 749	School	2	4	NA	5
	Teacher	92	216	NA	307
	Student	1,039	2,055	NA	3,094
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	1,024	598	457	2,079
	Teacher	10,055	5,815	2,561	18,431
	Student	70,758	52,713	19,156	142,627

Part IV - Basic GLS SASS total (139 schools)

School Size		Central City	Urban Fringe / Large Town	Rural / Small Town	Total
1 - 149	School	912	522	438	1,872
	Teacher	6,666	3,438	2,456	12,560
	Student	44,299	29,240	17,886	91,425
150 - 299	School	55	62	9	126
	Teacher	1,180	1,131	187	2,498
	Student	9,053	11,136	1,641	21,831
300 - 499	School	50	29	NA	79
	Teacher	2,070	1,174	NA	3,244
	Student	16,579	11,513	NA	28,091
500 - 749	School	1	2	NA	2
	Teacher	30	99	NA	129
	Student	338	941	NA	1,280
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	1,017	614	447	2,079
	Teacher	9,946	5,842	2,643	18,431
	Student	70,269	52,830	19,527	142,627

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 8.4 -- Nonsectarian Special Emphasis: Estimates by school size and community type, PSS and SASS compared in percent difference

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and original SASS totals					
1 - 149	School	-1.03	36.73	7.27	14.75
	Teacher	3.58	43.41	30.22	23.60
	Student	8.82	27.96	11.81	16.43
150 - 299	School	35.11	-6.19	74.22	26.65
	Teacher	27.77	13.70	78.32	31.67
	Student	45.95	6.23	76.57	36.47
300 - 499	School	-48.26	-29.22	NA	-29.27
	Teacher	-77.06	3.09	NA	-28.51
	Student	-41.16	-34.45	NA	-25.58
500 - 749	School	-53.42	-5.86	NA	-27.85
	Teacher	-34.11	-31.70	NA	-32.91
	Student	-54.82	-10.39	NA	-31.12
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total % diff from PSS	School	-0.04	32.08	11.66	14.28
	Teacher	-4.05	32.24	40.82	18.00
	Student	6.82	14.02	30.88	13.47
Percent difference from PSS and Olkin GLS SASS totals					
1 - 149	School	-19.16	26.75	-9.52	-0.07
	Teacher	-18.89	32.72	15.72	7.07
	Student	-8.17	17.92	-3.90	2.24
150 - 299	School	27.08	-3.55	73.46	23.56
	Teacher	15.75	6.52	76.81	23.35
	Student	39.87	6.14	75.93	33.40
300 - 499	School	-68.81	-71.23	NA	-55.70
	Teacher	-105.03	-68.07	NA	-71.33
	Student	-59.66	-83.91	NA	-53.54
500 - 749	School	15.98	-40.23	NA	-14.56
	Teacher	26.64	-74.69	NA	-23.67
	Student	15.31	-46.43	NA	-17.63
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total % diff from PSS	School	-17.25	22.02	-4.14	0.00
	Teacher	-24.34	13.72	28.95	0.00
	Student	-6.60	-0.34	19.22	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 8.4 – Nonsectarian Special Emphasis: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		Community Type: Central City	Community Type: Urban Fringe / Large Town	Community Type: Rural / Small Town	Total % diff from PSS
		Percent difference from PSS and Basic GLS SASS totals			
1 - 149	School	-18.30	24.83	-6.23	0.30
	Teacher	-18.73	26.89	15.25	4.95
	Student	-9.31	13.24	-2.68	0.24
150 - 299	School	19.53	-23.39	58.74	10.51
	Teacher	6.48	-7.95	63.30	11.38
	Student	32.94	-10.71	62.64	21.90
300 - 499	School	-57.58	-47.37	NA	-40.93
	Teacher	-90.75	-35.46	NA	-50.97
	Student	-48.98	-56.19	NA	-37.71
500 - 749	School	72.60	35.55	NA	52.74
	Teacher	76.11	19.97	NA	48.24
	Student	72.42	32.92	NA	51.34
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total % diff from PSS	School	-16.52	19.85	-1.80	0.00
	Teacher	-22.99	13.33	26.67	0.00
	Student	-5.86	-0.56	17.65	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

3.9 NONSECTARIAN SPECIAL EDUCATION TYPOLOGY

The Special Education typology represents almost the smallest type of private school. For example, in the 1993-94 Private School Survey, there were just an estimated 1,237 such schools or about 4.7% of the private school total for that year.

In table 9.1, SASS and PSS estimates are shown for schools, teachers, and students. Bringing these three SASS totals in line with the corresponding PSS ones is the exercise to be engaged in here. Notice that the SASS school totals are lower than the PSS (by about 3%); SASS also underestimated teachers and students relative to the PSS (by 8.3% and 16.6% respectively).

To set the stage for the calculations that follow, it might be worth looking at figure 9.1 which provides a scatterplot of student enrollment by number of teachers. The PSS sample schools are shown as black □'s. Overlaying the PSS data and shown as gray Δ's are the data from the 168 SASS sample cases.

As can be seen, the SASS scatter lies within that for PSS and along essentially the same axis. In fact, the slope of the student/teacher relationship is 4.0 for PSS and 4.1 for SASS -- virtually indistinguishable (Indeed, the least squares lines are touching over most of their length). While not directly comparable because of differences in sample designs, the R^2 values for the student/teacher ratios in both (unweighted) samples show analytically what can be observed from the graph, namely that the scatter is not very tightly bunched around the average teacher/student relationship. These values are $R^2 = .31$ (PSS) and $R^2 = .18$ (SASS).

Since these surveys are for the same year, the expected values for the weighted PSS and corresponding SASS quantities would be the same; hence an estimator that made them exactly equal might be an appropriate condition to impose.

In constructing these new GLS estimators, a decision was first made about which sample cases to use (see section 3.9.1); then the Olkin factors were calculated and applied to the original SASS weights, before carrying out the GLS adjustment (see section 3.9.2). The results of the basic GLS were also obtained (section 3.9.3). Operational considerations are covered next and comparisons made to the original SASS weighted sample and to the Olkin GLS and basic GLS versions (section 3.9.4). An independent assessment (section 3.9.5) concludes the discussion.

3.9.1 Determining Outliers. -- Before applying a GLS adjustment, the first step taken was to look closely at the scatter in figure 9.1 -- to see if any SASS or PSS outliers should be excluded from the GLS reweighting and handled in another way. See Section 4.4 for more discussion of this.

Table 9.1 – Nonsectarian Special Education: Weighted schools totals before excluding outliers (Based on 1,086 PSS and 168 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	1,237	1,274	-37
Teachers	13,695	14,844	-1,149
Students	74,087	86,356	-12,269

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

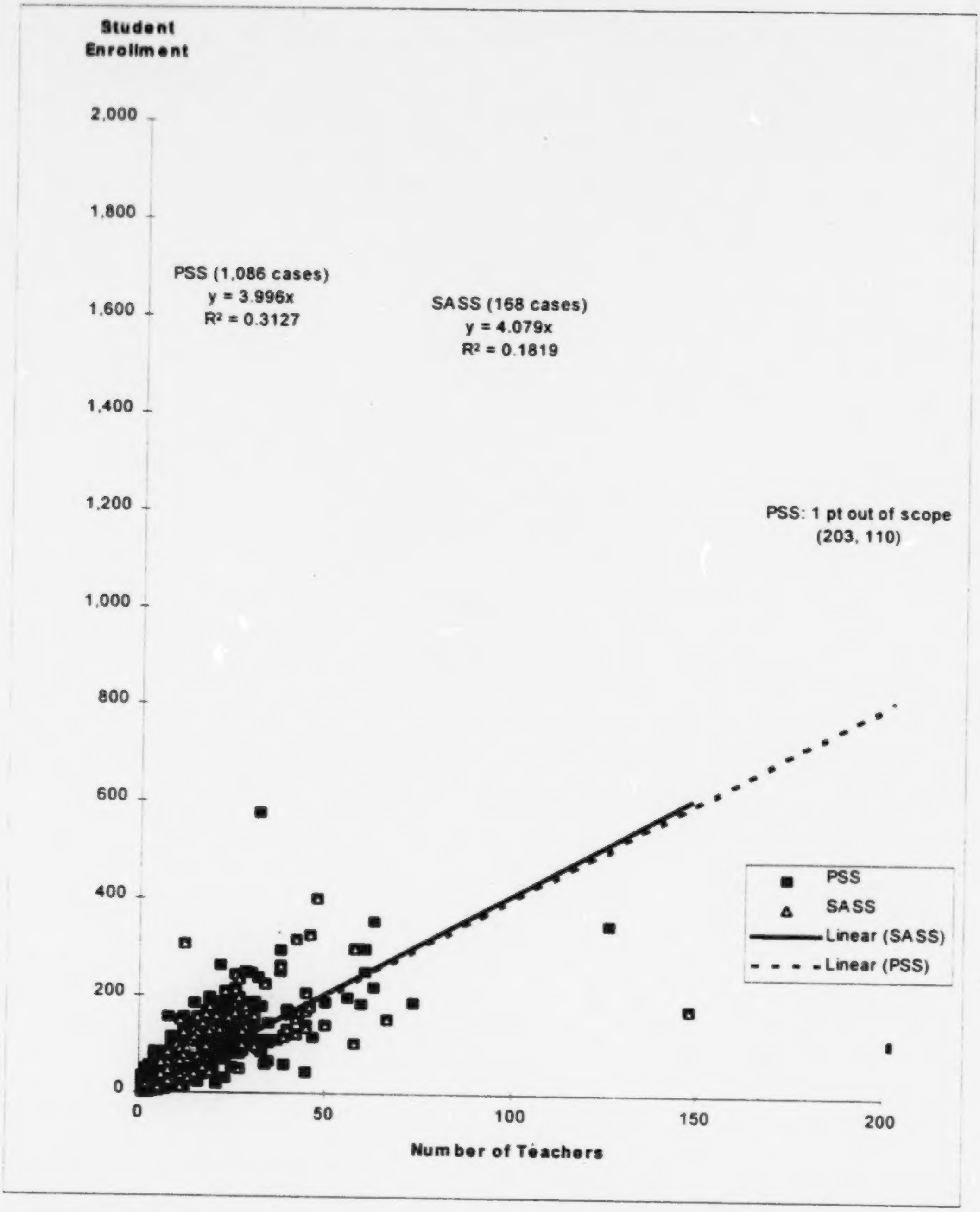
Table 9.2 – Nonsectarian Special Education: Weighted schools totals after excluding outliers (Based on 1,079 PSS and 165 SASS sample schools)

Variable	PSS	SASS	Difference
Schools	1,229	1,268	-39
Teachers	13,006	14,487	-1,481
Students	71,592	84,659	-13,067

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 9.1 -- Nonsectarian Special Education: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

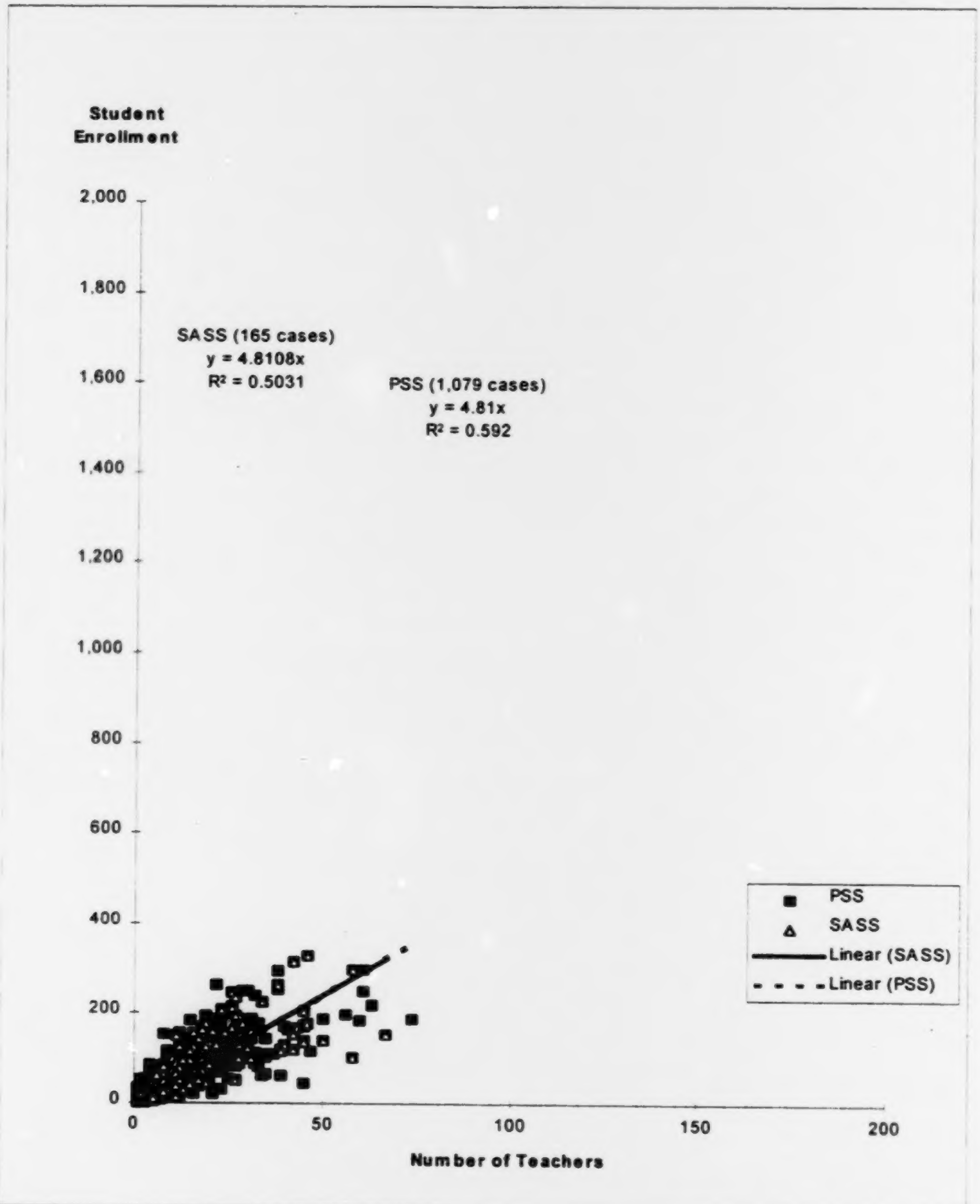
(before removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Figure 9.2 – Nonsectarian Special Education: Student versus teacher unweighted sample totals for PSS and SASS combined in 1993-94

(after removing outliers)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

BEST COPY AVAILABLE

For the Special Education typology, simple visual inspection seemed sufficient, resulting in a reduced PSS sample(from 1086 to 1079 cases)and a correspondingly reduced SASS sample (from 168 to 165 cases). Figure 9.2 is the plot of the remaining cases. Notice that the student/teacher relationships have changed somewhat from those in figure 9.1; additionally, the scatter of points in both samples is visually much tighter.(The R^2 values have both increased sizably too.)

3.9.2 Olkin GLS Procedure Employed. -- To carry out the Olkin GLS weighting the equation

$$\underline{\boldsymbol{\lambda}} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

needed to be solved, as described in Section 2; but only after adjusting the original SASS weights within school size classes. The starting point was the new typology totals for PSS and SASS. These are shown in table 9.2 below.

To carry out the Olkin GLS, the schools were placed into two school size classes (specifically under 150, 150 and above).

After the Olkin adjustment, the difference between the PSS and SASS estimates had shrunk (considerably in overall absolute value) to

$$\underline{\mathbf{d}} = \begin{matrix} -26 \\ 346 \\ -84 \end{matrix}$$

The matrix **M** was obtained by tabulating the 1993-94 SASS file for the eligible Special Education schools in the SASS sample. The values are

$$\begin{matrix} 165 & 2638 & 14515 \\ 2638 & 69092 & 332386 \\ 14515 & 332386 & 1917245 \end{matrix}$$

Solving for $\underline{\boldsymbol{\lambda}}$ yields

$$\underline{\boldsymbol{\lambda}}' = (-.53570, +0.03714, -0.00243)$$

and the Olkin GLS weights are of the form

$$u_i = w_i - 0.53570 + 0.03714t_i - 0.00243s_i$$

Notice that all the original weights are lowered (by about .5); and, then, depending on the teacher and student counts in the sampled school, they may be increased or lowered (usually they would not be lowered further except for the very largest schools). These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts are so small. One final comment, while the values for $\underline{\lambda}$ are only shown to six significant digits, the calculations have been carried out in double precision.

3.9.3 Basic GLS Procedure Employed. -- To carry out the basic GLS weighting the equation

$$\underline{\lambda} = \mathbf{M}^{-1}\underline{\mathbf{d}}$$

again needs to be solved. It is immediate from table 9.2 above that $\underline{d}$ for the basic GLS would be

$$\underline{\mathbf{d}} = \begin{matrix} -39 \\ -1481 \\ -13067 \end{matrix}$$

The matrix **M** is again obtained by tabulating the eligible portion of the 1993-94 SASS file for the eligible Special Education schools in the SASS sample. Because the Olkin and Basic GLS employ just the unweighted sample to calculate **M**, it is the same for both (and hence not shown).

Solving for $\underline{\lambda}$ yields this time

$$\underline{\lambda}' = (+.96094, +0.05823, -0.024185)$$

and the basic GLS weights are of the form

$$u_i = w_i + 0.96094 + 0.05823t_i - 0.024185s_i$$

Notice that all the original weights are raised this time (where for the Olkin GLS they were lowered); and, then, depending on the teacher and student counts in the sampled school, they may be increased again or lowered (usually they would not be "lowered," except for the very largest schools). These additional school-by-school adjustments do not appear to be too drastic -- given that the coefficients on the teacher and student counts are so small. They are the same size or somewhat larger in absolute value, though, than for the Olkin adjustment -- a pattern that was expected (and which turns out to be generally true overall).

3.9.4 Operational Characteristics. -- To examine the Basic and Olkin GLS reweighting done, several "diagnostics" will be looked at. One statistic that may merit immediate attention is what happened to the spread in the weights themselves. Did the spread grow larger or smaller?

Figure 9.3 provides this information in its upper panels, which compare the original and two GLS adjustments. The Basic GLS weight has a smaller spread than does the original SASS weight. This follows by noting that the original (or y) weights in the upper panel are slightly smaller than the GLS (or x) weights (since the equation which fits them is $y = .9949x$). For the Olkin GLS, the variability in the weights is somewhat smaller still than in the original SASS (with the equation relating them being of the form $y = .9598x$).

While the overall differences in scale between the weights appear unimportant, the scatter for the Olkin GLS unaccountably shows a distinct break between the original data and the final Olkin weights for the smallest schools.

The R^2 values shown in the upper panel in figures 9.3 might be commented on too. Despite the appearance of the scatter itself, very little evidence exists to indicate that either of the reweighting approaches altered the original weights in any major way. The R^2 values are both at or above 0.97 and most of the points lie close to the 45 degree line. The problem of negative weights did not arise for the Olkin GLS, although there were ten schools with weights smaller than one. For the Basic GLS, the results were not quite as good. Negative weights occurred for 8 schools, and there were 19 more cases with weights less than one.

A brief comment might be made on the differences between the Basic and Olkin GLS weights. Figure 9.3 will continue to be our source. This time, though, look at the panel in the second row of the scatterplot matrix where the Basic and Olkin GLS are graphed. The Basic GLS (or y) values are smaller than the Olkin GLS (or x) values; but ever so slightly. There is virtually no overall difference in the weights -- as evidenced by an R^2 of .95 between the two methods. The plotted points do indicate some departures though, as noted earlier, among a handful of schools.

3.9.5 Independent Assessments. -- The ingredients used here for an independent assessment of the GLS adjustment of the Special Education Typology are available in tables 9.3 and 9.4, plus figure 9.4:

-- Table 9.3 is a cross-tabulation of PSS and SASS by community type and school size. Within each combination of these, PSS totals for schools, teachers, and students can be compared with the corresponding original SASS, Olkin, and Basic GLS estimates respectively.

-- Table 9.4 is based on table 9.3 but focuses directly on percentage differences between the three SASS estimates and PSS.

-- Figure 9.4, which provides the corresponding graphical summaries, is where the discussion is mainly concentrated. Note this graph is on a log scale.

One place to begin an assessment is by determining the degree to which the various reweighted SASS results agreed to the PSS by size of school. After some looking at the total columns in tables 9.3 and especially 9.4, it is clear that the Olkin GLS, while far from uniformly better, certainly does the best on the average. Of the eighteen overall comparisons by school size, the Olkin GLS weighted SASS is closer to the PSS than the original weighted SASS, 11/18 times; and closer than the Basic GLS also in 11/18 comparisons. Thus, in over half of the cases, the Olkin method is to be favored.

The results by community type are not very good for any of the estimators. This, while disappointing, might have been expected since none of the approaches looked at community type (and the typology sample size is small).

In figure 9.4, the plots may provide an overall sense of what is happening. The top panel in that figure is the place to begin. Notice first that the original SASS weighted file was on the average already fairly close to the PSS, since the best fit regression equation which connects the various estimates is $y = .973x$. There is a little roughness around this average, however, as displayed visually and summarized by the R^2 value which is equal to $R^2 = .9617$.

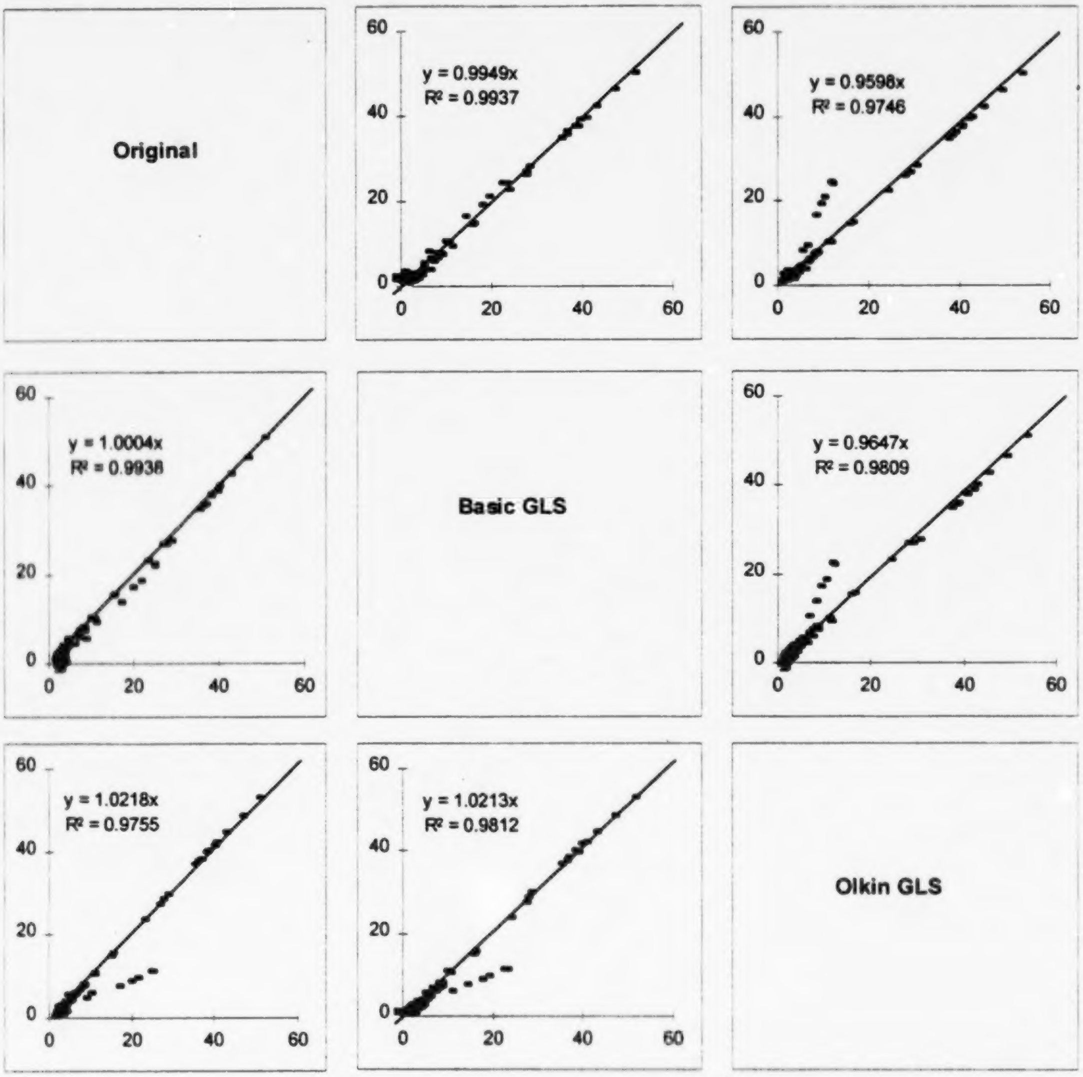
Also in figure 9.4, it can be seen that for the Basic GLS weighted SASS sample the regression fit is similar, yielding the relationship $y = 1.0065x$. Again, the average results for this method remain good. Considerably more roughness is exhibited, though, around the average as evidenced by the smaller R^2 value in this case ($R^2 = .834$).

Finally, in figure 9.4, it can be seen that for the Olkin GLS weighted SASS sample the regression fit yields the relationship $y = 1.0099x$. The average results for this method are close to the best, with an R^2 value in this case of $R^2 = .9745$.

What can be concluded about this typology? Neither the Basic nor the Olkin GLS methods seem in any way inferior overall to the original SASS weighted file. To their credit, they both hit the overall PSS school, teacher, and student totals exactly. The Olkin GLS method, furthermore and not surprisingly, does as well or better than the other two when estimates are looked at by school size. The Basic GLS method has negative weights and does not fair as well when external comparisons are made.

In the summary and recommendations section, some further comments will be made about how the Olkin GLS might be improved further, leading to still better results.

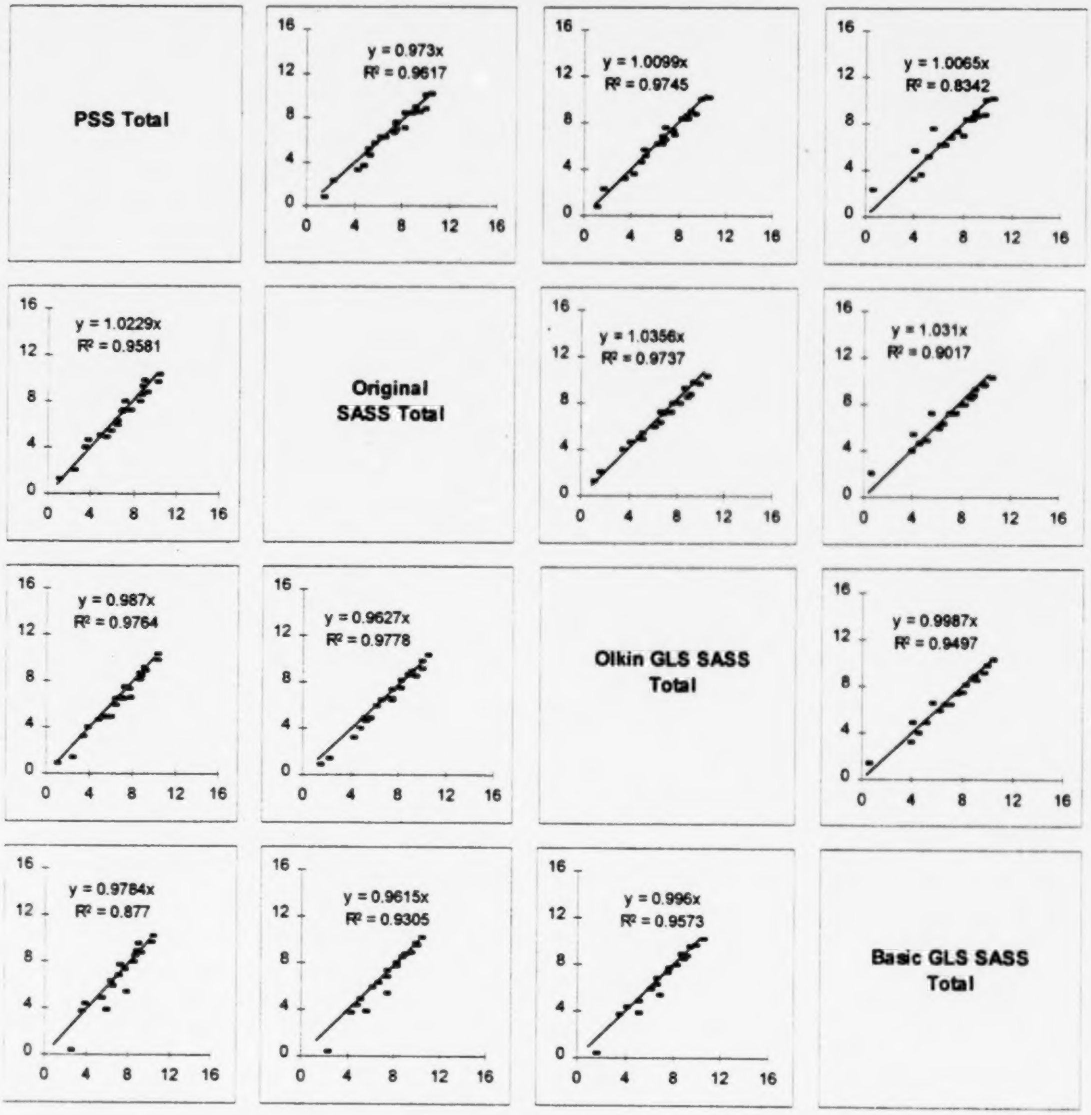
Figure 9.3 – Nonsectarian Special Education: Scatterplot matrix comparing original, basic, and Olkin GLS weights



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

BEST COPY AVAILABLE

Figure 9.4 – Nonsectarian Special Education: Scatterplot matrix comparing estimated PSS, original, Olkin GLS, and Basic GLS SASS totals by school size and community type from Table 9.3 (in log scale)



SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94.

Table 9.3 -- Nonsectarian Special Education: Estimates by school size and community type, PSS and SASS compared

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part I - PSS total (1,079 schools)					
1 - 149	School	496	482	177	1,155
	Teacher	4,267	4,675	1,624	10,566
	Student	23,409	25,424	8,159	56,992
150 - 299	School	26	36	10	73
	Teacher	955	1,090	296	2,340
	Student	5,331	6,613	1,928	13,872
300 - 499	School	NA	2	NA	2
	Teacher	NA	100	NA	100
	Student	NA	727	NA	727
500 - 749	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	522	521	187	1,229
	Teacher	5,222	5,864	1,920	13,006
	Student	28,740	32,765	10,087	71,592
Part II - Original SASS total (165 schools)					
1 - 149	School	382	583	137	1,101
	Teacher	3,044	5,325	1,387	9,757
	Student	16,765	29,495	6,872	53,132
150 - 299	School	55	101	8	164
	Teacher	1,316	3,032	223	4,570
	Student	10,575	18,406	1,384	30,365
300 - 499	School	NA	4	NA	4
	Teacher	NA	160	NA	160
	Student	NA	1,161	NA	1,161
500 - 749	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	436	688	144	1,268
	Teacher	4,360	8,517	1,609	14,487
	Student	27,340	49,062	8,256	84,659

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 9.3 -- Nonsectarian Special Education: Estimates by school size and community type, PSS and SASS compared (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Part III - Olkin GLS SASS total (165 schools)					
1 - 149	School	394	608	142	1,144
	Teacher	3,229	5,642	1,540	10,410
	Student	17,386	30,838	7,220	55,444
150 - 299	School	26	53	4	83
	Teacher	616	1,742	134	2,492
	Student	4,916	9,717	758	15,391
300 - 499	School	NA	2	NA	2
	Teacher	NA	104	NA	104
	Student	NA	756	NA	756
500 - 749	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	419	664	146	1,229
	Teacher	3,845	7,488	1,673	13,006
	Student	22,302	41,311	7,978	71,591
Part IV - Basic GLS SASS total (165 schools)					
1 - 149	School	388	583	142	1,113
	Teacher	3,065	5,251	1,526	9,843
	Student	16,410	28,375	6,729	51,515
150 - 299	School	40	79	1	121
	Teacher	936	2,394	46	3,375
	Student	7,415	13,997	214	21,627
300 - 499	School	NA	-5	NA	-5
	Teacher	NA	-212	NA	-212
	Student	NA	-1,550	NA	-1,550
500 - 749	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total	School	428	658	143	1,229
	Teacher	4,001	7,433	1,572	13,006
	Student	23,826	40,822	6,944	71,592

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

Table 9.4 – Nonsectarian Special Education: Estimates by school size and community type, PSS and SASS compared in percent difference

School Size		Community Type			Total % diff from PSS
		Central City	Urban Fringe / Large Town	Rural / Small Town	
Percent difference from PSS and original SASS totals					
1 - 149	School	23.03	-21.00	22.84	4.63
	Teacher	28.65	-13.91	14.60	7.66
	Student	28.38	-16.01	15.77	6.77
150 - 299	School	-107.77	-176.78	20.29	-125.34
	Teacher	-37.75	-178.25	24.69	-95.27
	Student	-98.36	-178.32	28.20	-118.89
300 - 499	School	NA	-59.03	NA	-59.03
	Teacher	NA	-60.66	NA	-60.66
	Student	NA	-59.58	NA	-59.58
500 - 749	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total % diff from PSS	School	16.41	-32.08	22.71	-3.16
	Teacher	16.51	-45.24	16.16	-11.39
	Student	4.87	-49.74	18.15	-18.25
Percent difference from PSS and Olkin GLS SASS totals					
1 - 149	School	20.54	-26.28	19.74	0.88
	Teacher	24.33	-20.67	5.19	1.48
	Student	25.73	-21.30	11.51	2.72
150 - 299	School	3.07	-44.94	56.75	-13.89
	Teacher	35.51	-59.89	54.83	-6.47
	Student	7.79	-46.92	60.68	-10.95
300 - 499	School	NA	-3.52	NA	-3.52
	Teacher	NA	-4.66	NA	-4.66
	Student	NA	-3.97	NA	-3.97
500 - 749	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total % diff from PSS	School	19.66	-27.48	21.67	0.00
	Teacher	26.38	-27.69	12.83	0.00
	Student	22.40	-26.09	20.91	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

BEST COPY AVAILABLE

Table 9.4 – Nonsectarian Special Education: Estimates by school size and community type, PSS and SASS compared in percent difference (cont'd)

School Size		Community Type			Total
		Central City	Urban Fringe / Large Town	Rural / Small Town	% diff from PSS
		Percent difference from PSS and Basic GLS SASS totals			
1 - 149	School	21.72	-21.09	19.81	3.56
	Teacher	28.16	-12.32	6.01	6.85
	Student	29.90	-11.61	17.53	9.61
150 - 299	School	-51.67	-117.55	85.58	-66.46
	Teacher	2.04	-119.68	84.45	-44.22
	Student	-39.10	-111.64	88.87	-55.90
300 - 499	School	NA	313.66	NA	313.66
	Teacher	NA	312.41	NA	312.41
	Student	NA	313.03	NA	313.03
500 - 749	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
750 +	School	NA	NA	NA	NA
	Teacher	NA	NA	NA	NA
	Student	NA	NA	NA	NA
Total % diff from PSS	School	18.01	-26.39	23.23	0.00
	Teacher	23.38	-26.74	18.09	0.00
	Student	17.10	-24.59	31.16	0.00

SOURCE: U.S. Department of Education, NCES, Private School of Schools and Staffing Survey, 1993-94, Private School Surveys, 1993-94. "NA" means that there were no observations in the table cell.

4. SUMMARY AND RECOMMENDATIONS

In this section, there are some reflections on the experience, just documented, of employing variants of a Generalized Least Squares (GLS) approach to two of NCES's most important surveys. We begin with a restatement of the problem posed and the basic approach taken (Subsection 4.1). A summary of the results obtained follows in some detail (Subsection 4.2). Next there are two subsections which make recommendations: Subsection 4.3 focuses on further efforts at GLS and GLS-like estimation methods that might be tried. There are many interesting challenges here and the opportunities for improvements appear considerable. Subsection 4.4 takes a different tack and looks at a method called "mass imputation" which could also deserve study in a SASS setting.

4.1 PROBLEM RESTATEMENT AND BASIC APPROACH

This report has provided empirical results of attempts to achieve consistency between the 1993-94 Private School Survey (PSS) and the Private School Component of the 1993-94 Schools and Staffing Survey (SASS). As the PSS is the basis for the SASS sampling frame, the PSS results, on the whole, are likely to be the more accurate. Under these circumstances, it made sense to explore whether the introduction of 1993-94 PSS totals into the 1993-94 SASS might lead to improvements.

Traditional post-stratification methods exist to employ auxiliary information at the estimation stage in surveys. These, however, cannot be applied to SASS without modification, since consistency was sought simultaneously in the numbers of schools, teachers, and students from these two sources. This led us to employ various forms of Generalized Least Squares (GLS) estimation to reweight the 1993-94 SASS. Two variants were looked at in the main: a Basic GLS procedure (Burton 1989) and a method we have dubbed the Olkin GLS because it is a variant suggested by Olkin (1958) which arose originally in a different context.

As we have seen, for the private school population nine typologies exist which differentiate schools by whether they are Catholic, Other Religious, or Nonsectarian. There is then a further subdivision into three additional groups: Catholic (Diocesan, Diocesan, or Private); Other Religious (Affiliated with a conservative Christian school association, Affiliated with national denomination or other religious school association, or Unaffiliated); and Nonsectarian (Regular programs, Special emphasis, or Special education). For each of these nine typologies, we separately attempted to achieve intersurvey consistency. Sometimes this was straightforward; sometimes extremely difficult.

The complex nature of the PSS and SASS sample designs was considered in the approach taken. Operational problems were documented; and independent comparisons were made to PSS school size and community type information that was not used directly in the reweighting. Measures of benefit and harm could be developed because of the comparisons possible. Extensive tabular, graphical, and analytic material have been looked at in making the assessments required.

The summary of our results that follows is thus grounded in an extensive body of empirical evidence for the 1993-94 SASS and PSS. The present work also builds directly on an earlier pilot effort involving the 1991-92 PSS and the 1990-91 SASS which in many ways was almost as extensive (See Li and Scheuren 1995).

4.2 SUMMARY OF EMPIRICAL RESULTS

The summary given here is divided into three parts: a review of the Olkin GLS, a assessment of the Basic GLS and a few concluding overall remarks.

4.2.1 Olkin GLS Summary. -- As can be seen in table A, the operational assessment of the Olkin GLS adjustment to SASS was judged to be good to excellent. In only one case, that for the Other Unaffiliated typology was the evidence unclear. We consider this typology unclear because the Olkin GLS did not work without a considerable amount of ad hoc tinkering (See Subsection 3.6).

Table A.-- Olkin GLS Comparisons to Original Weighted SASS Data, By Typology

SASS Typology	Operational Assessment	Independent Assessment
Catholic Parochial	excellent	good
Catholic Diocesan	excellent	fair
Catholic Private	excellent	good
Conservative Christian	good	fair
Other Affiliated	excellent	good
Other Unaffiliated	unclear	good
Non-sectarian Regular	good	good
Non-sectarian Special Emphasis	good	good
Non-sectarian Special Education	good	fair

Based on the independent assessment by community type and school size, the Olkin GLS seemed to do no apparent harm and may have even been of benefit -- beyond the basic consistency achieved with PSS. The comparisons made are to the original SASS weighted data.

The admittedly subjective conventions employed in table A were devised to separate typologies by level of perceived difficulty or benefit. Operationally

> -- typologies where a simple visual inspection was all that was needed to remove outliers are labelled "excellent" in the operational assessment column.
>
> -- typologies labelled "good" were ones where an analytic (potentially iterative) process was required to identify SASS cases that might best be treated by imputation to similar PSS cases rather than being

reweighted.

-- only the Other Unaffiliated typology is labelled "unclear." This was done because, as noted earlier, constructing the Olkin GLS weights was enormously difficult and required great patience and persistence. (Parenthetically it may, also, have been the most instructive in terms of learning more about how to employ the GLS.)

The independent assessment column was never coded "excellent" because, especially by community type, the Olkin GLS was never best overall. Regularly, it did a "good" job, usually by school size, but even here the performance was less than hoped. In three cases, the Olkin GLS was judged only "fair." These were instances where very mixed results were achieved: some estimates much improved, others quite negatively impacted.

4.2.2 Basic GLS Summary. -- Using the results of table B below, a summary of the Basic GLS is given. Looking at the independent assessment by community type and school size, the Basic GLS seemed to do little apparent harm and may have even been of benefit, beyond the basic consistency achieved with PSS.

The subjective coding of the results in table B is based on a lower set of expectations for the Basic GLS than for the Olkin GLS. Operationally

-- typologies identified as "good" were ones with no negative and no more than a few small weights (i.e., weights less than one).

-- the typology identified as "fair' was so labelled because, while no negative weights arose, there were a great many small weights (under one).

-- typologies labelled "poor" were those having negative weights.

For the independent assessment column, a somewhat more liberal interpretation of "good" and "fair" are given than was the case in discussing the Olkin GLS. Frankly, based on our earlier pilot, there was a lower expectation. But, relative to this expectation, while the results (as for the Olkin GLS) were never "excellent," they were often surprisingly good.

By "good" we meant that the Basic GLS reweighted SASS tracked the PSS quite well on an overall basis, as judged by the regression results. To rate only a "fair," the Basic GLS had to perform less well in the regressions than did the original SASS -- notably having an R^2 value indicating a greater degree of roughness in the PSS relationship.

Table B. -- Basic GLS Comparisons to Original SASS Weighted Data, By Typology

SASS Typology	Operational Assessment	Independent Assessment
Catholic Parochial	good	good
Catholic Diocesan	poor	good
Catholic Private	good	good
Conservative Christian	poor	fair
Other Affiliated	fair	good
Other Unaffiliated	poor	good
Non-sectarian Regular	poor	good
Non-sectarian Special Emphasis	good	good
Non-sectarian Special Education	poor	fair

4.2.3 Overall Summary. -- It may be instructive to look back on the summary made in the pilot work done about a year ago (Li and Scheuren 1995). This has been done below. The backward glance helps frame what was learned and what was confirmed. As will be seen, there is a sense of real progress in some areas and grounds for optimism in others:

(1) In the initial attempt at GLS, it appeared that in under half of the typologies no or minimal harm was done. Moreover, in some of the other typologies, the GLS adjustment may have caused severe problems with many large schools having sizable negative weights.

> **With the new Olkin GLS approach taken, this is no longer true.**

(2) Earlier, a closer look at the cases where no apparent harm occurred revealed that these were situations where almost no adjustment was needed to begin with. This made it reasonable to assume that, for the application of the Basic GLS to the 1993-94 SASS, fair to good results should be expected. After all, unlike in the test done for this report, both PSS and SASS were collected for the same school year.

> **This expectation has largely been borne out, although the negative weight problem persisted for some typologies for the Basic GLS.**

(3) Only a few experiments were tried in the pilot to handle negative weights. These have been continued here. There are, however, a number of methods to dampen the effects of negative weights or even eliminate them that seem practical (e.g., Huang 1978).

> **Since then still others have become known (e.g., Brewer 1994). More research is needed in this area as discussed in section 4.3 below.**

(4) The experience gained in compiling operational statistics on the workings of the GLS adjustments was instructive. Only some of these variables may need to be tracked, though, since they so frequently gave the same "bottom line."

> **For the current report, the statistics used to track operational performance were cut back successfully. However, there was in one case an addition to the performance measures -- a cumulative weight distribution comparison.**

(5) Independently assessing the GLS reweighting proved particularly instructive (in the pilot) and allowed both bias and variance effects to be looked at -- i.e., among the most natural benefit and harm measures. Using this additional information in the adjustment would perhaps have improved results; but retaining at least some outside data for evaluation seemed essential too.

> **The employment in the current report of school size information in the adjustment was a result of this observation. Community type data was added too, so as to retain a completely independent comparison.**

(6) The use of a modified GLS reweighting, even when it is beneficial, does not make much of a positive difference beyond achieving consistency with PSS. Other methods, done separately or in combination with GLS, appear needed in order to take full advantage of the opportunity that having PSS and SASS fielded for the same year offers.

> **Here the Olkin GLS was introduced as a partial answer plus the separate treatment, through imputation of the largest schools. Still the positive benefits were often disappointing.**

More was hoped in last year's report than has been delivered. Part of the problem is that expectations were misplaced. After all, why should introducing just three totals from PSS make a big improvement into SASS. Conversely, why should such a seemingly small change sometimes be so hard?

4.3 RECOMMENDATIONS FOR FURTHER REWEIGHTING RESEARCH

Some recommendations are implied in the summary discussion given above for future research on reweighting SASS. As will be seen, many of these are quite concrete -- others are more conjectural or of a basic research nature.

An overriding concern is that, even with what might be characterized as considerable experience, it is hard to predict when the GLS estimator will perform well and when it will not. Theoretical and computational developments continue (e.g., Kott 1996) with GLS estimators, so any suggestions made here should be taken as no more than tentative. Practice has been expanding considerably as well (e.g., Jayasuriya and Valliant 1995).

Recommendation 1

Except for special circumstances, the Basic GLS approach should not be relied on alone in SASS.

Based on our experiences, the Basic GLS is expected to work well only in typologies where the SASS and PSS samples are close to begin with and none of the original SASS weights were small. If all the totals are off by about the same percentage, then GLS degenerates to essentially a univariate post-stratification procedure and the usual "rules of thumb" apply. In univariate raking ratio estimation, one such "rule" is to have about 20 to 25 observations per dimension being constrained (e.g., Oh and Scheuren 1978). In the SASS application, this means that we would need 3X20 = 60 cases in each group for which GLS reweighting was to be attempted. Naturally, if the three totals in SASS differ from PSS in quite different degrees and especially if they differ in different directions, then a more conservative "rule" is plausible -- albeit, as yet, unproven.

Recommendation 2

Further work might profitably be done on the Olkin GLS, especially on the theoretical relationship which would appear to connect Olkin's original ideas from 1958 with the new work continuing on GLS.

Operationally, the Olkin GLS could have been more aggressively pursued by using smaller sized groups (say, with only 60 to 100 or so schools each, as eventually occurred for the Other Unaffiliated typology). Clearly, too, the groups should have been better chosen so as to conform with the publication plans for SASS. Implementing this recommendation might not be difficult given the approach taken here. Theoretical work is needed on GLS, especially in connecting it better to the insights Olkin has nearly 40 years ago; however, this might not be a priority for NCES. A wait and see approach could work quite well, since so much is being done elsewhere.

Recommendation 3

The Olkin adjustment might be used alone, without the GLS.

This suggestion is made because the process is quite easy and for 1993-94 no harm was done in most cases. This takes advantage of the univariate portion of the GLS estimation, which in many cases was of benefit and very low cost. Certainly no negative weights are possible. The potential for benefits are considerable too, especially if done in moderate sized groupings (as suggested in Recommendation 2 above).

Recommendation 4

> ***Methods for variance estimation need exploration. While the general GLS approach is well covered in the literature, an efficient method has to be programmed and tested in the SASS environment.***

Of course, concerns exist, too, about the impact on variance and variance estimation of the various ad hoc adaptations needed to keep the weights reasonable. We have found for the private school population and the existing SASS design that negative weights occur frequently enough to question whether the asymptotic variance formulas can be used with safety. A bootstrapping approach makes sense here, if the computational costs can be borne.

Recommendation 5

> ***Some improvements in SASS and PSS processing may be a consequence of the study of GLS applications. One of those that has arisen so far is the clear possibility (see Holt et al 1994) that SASS edit checking could be enhanced if GLS estimation is attempted.***

A subtler concern is the treatment in SASS of the very largest schools, when these become nonrespondents. Here perhaps an imputation rather than a reweighting approach may be preferred -- using, say, the PSS data as a starting point (e.g., Kovar et al 1994, Scheuren et al 1996).

Among schools above a given size imputation might have more benefit in reducing SASS mean square error than GLS. This was assumed by us as the way we would treat "outliers." Much more could be done, though, especially on where to place the boundary between where weights are used and where an imputation is employed. This recommendation is expanded on further in the next subsection where "mass imputation" is covered.

Recommendation 6

> ***There is a real need to look at the sample design in SASS and see if it is partially responsible for the performance problems that GLS, in its various forms has had.***

At present, SASS is a stratified sample where a function of the number of teachers is the principle design variable. Schools that are far from the average student/teacher ratio in their typology can contribute greatly to the variances for student charactersitics since enrollment size is not directly controlled for in the design. As has been seen, it is these schools that are more likely to receive negative weights. A design, where both teachers and students are stratifiers could reduce the variability on the student enrollment enough so that negative weights from this cause became infrequent, especially when combined with some of the other suggestions being made.

Another sample design issue has to do with the continuing use of an area frame for SASS. Instead of having both an area frame in SASS and PSS, there might only be one in PSS. To obtain the SASS estimate, the starting point might be the list frame portion of the PSS adjusted for undercoverage (e.g., Causey, Bailey, and Hoy 1996); then projected so the PSS totals could be used as controls for SASS. In any event, discontinuing the area component of SASS seems something to look at (as discussed in Scheuren et al 1996). If it is not discontinued, then among the issues to consider is whether to jointly determine the estimate from PSS and SASS for the area portion of the private school universe. A combined area estimate might be a worthwhile improvement for both surveys.

Recommendation 7

> ***There is a real need to explore other adjustments to SASS so as to capitalize on the richness of the companion PSS, whether or not fielded for the same year.***

The use of alternative GLS estimators as in Deville et al (1993) could warrant examination. This is not seen as likely to improve much on the modified GLS approach suggested by Burton(1989); but, especially in combination with other ideas listed above, could be tested. (See Kaufman and Scheuren 1996 for a number of other estimation possibilities).

4.4 RECOMMENDATION FOR MASS IMPUTATION RESEARCH

In this subsection, the conduct of mass imputation research is examined in a SASS context as a possible alternative to some form of GLS. To begin the discussion, it might be worth providing some background on how our research on intersurvey consistency led to the notion that a "mass imputation" approach might be worth considering. The literature on mass imputation is then summarized -- with the promised recommendations for SASS and PSS and some "What Nexts" concluding the presentation.

4.4.1 Background and Definition . -- In this report, as part of our implementation of GLS, we have set aside a small number of observations that were thought to be "outliers". This was done to reduce the chance of getting negative weights or weights less than one. The number of outliers was small for each typology and confined almost entirely to the very largest schools.

Implicit in setting aside these SASS schools was the notion that we would impute them to the PSS rather than try to reweight them. To start this process we would give the SASS outliers their PSS weights. Essentially, except for nonresponse and coverage adjustments, this means they would just be self-representing, since their corresponding PSS weights were only slightly larger than one. The remaining PSS outliers could then be added to SASS after first imputing to them non-PSS data from a similar SASS observation. Employing some form of the well known "Hot Deck" might be one way to do this, for example. Adjustments might be needed to make the imputed SASS data consistent with the existing PSS information. The PSS information, would, however, not be altered.

Based on our 1994 research (Li and Scheuren 1995) we were convinced that a pure GLS reweighting strategy would not work. This was what lead us to suggest a mixed strategy where reweighting was still employed but perhaps an imputation approach was used for a handful of "outliers."

At some point, the question was asked: Why not do more than just a few imputations? In fact why not impute the entire SASS file to the PSS, in order to take full advantage of the opportunity that having PSS and SASS fielded for the same year offered? In other words, why not do "mass imputation"?

Mass imputation, now roughly 20 years old (Colledge et al. 1978), is really quite straightforward in concept. The technique imputes records from a survey back to the sampling frame; and, in a sense, operates in making estimates as if there had been a census. Mass imputation of sample survey data to a complete population file has been shown to work in some Canadian applications (e.g., Whitridge, Bureau and Kovar 1990; Kovar and Whitridge 1995). Moreover, when efficiently done, the costs of mass imputation appear only moderately larger than weighting.

Historically, the concern was at the analysis stage. For mass imputation to make sense, cheap computing is needed because the whole population has to be processed. Given this last observation, it is not surprising that the Canadians, with a population about 1/10th that of the U.S., were pioneers in this method.

4.4.2 Mass Imputation and Reweighting. -- In Kovar and Whitridge (1995), there is an excellent discussion of mass imputation. Among other things, they comment on the parallels that can exist between weighting and imputation. They call attention to the work of Folsom (1981) in this connection. Oh and Scheuren (1983) may be another useful reference. Evidence that imputation model sensitivity can be a serious problem exists, as they point out -- citing Cox and Cohen (1985), among others.

While the asymptotic properties of GLS and GLS-like estimators are known to be attractive, as has been seen, their finite sampling properties are not necessarily desirable. Could a mass imputation approach share the same asymptotic properties as GLS and behave better in small samples? Certainly operational concerns with GLS procedures about small or negative weights disappear. Exactly how difficult the mass imputation procedure would be to implement in SASS is unknown; but the challenge does not seem to be daunting.

The approach would be to give all the SASS cases their PSS weights (as was done just for outliers earlier), then for PSS cases not in SASS, a SASS case similar to it would be used to impute the SASS information. At the end, the PSS weights would remain unchanged but every PSS record would have appended to it either its own SASS data or imputed SASS data taken from another case. No reweighting would be needed. Because all the PSS data (including the PSS weights) would be used unaltered, the PSS totals (for schools, teachers, and students) would be "hit" exactly -- "solving" the problem by imputation that we originally set out to resolve by the GLS reweighting.

Of course, it should be noted, that, as with GLS-like estimators, the effect on SASS estimates not also available from PSS is unknown (and potentially could be harmful in some cases). Additionally, difficulties exist in calculating variances and covariances when using mass imputation. In Clogg et al (1991), mass imputation was employed with variances being estimated based on the theory underlying multiple imputation (Rubin 1987, 1996). For Hinkins and Scheuren (1986), where mass imputation was applied, variances were not calculated but a multiple imputation approach to their estimation was advocated. In another application, by Wong and Ho (1991), bootstrapping was employed successfully to calculate variances. We think a form of bootstrapping might be the best approach in SASS. The paper by Kaufman (1996) presents related work.

4.4.3 A Specific Proposal. -- An experiment attempting mass imputation in SASS definitely needs consideration. Suppose that mass imputation were to be conducted as part of an overall change in SASS estimation. How would it be done? Suppose, for the sake of discussion, that PSS and SASS were both conducted in the same year, as was true for 1993-94. What would the steps be? We will sketch these broadly. (See Kaufman and Scheuren 1996 for more details.)

Take a specific typology, "Other Religious Unaffiliated" Schools. As was seen earlier, there were 329 schools in the SASS sample with this designation. In the corresponding PSS for the same period, there were 3,193 such schools. The original SASS estimate of students in other religious unaffiliated schools was 462,934. From PSS, the estimate was 37,578 smaller-- at 425,356 students. Figure 6.5, discussed previously, compares cumulative PSS and SASS weighted survey observations by student enrollment for six school sizes separately. An Olkin GLS reweighting approach was taken to this problem to "solve it." However, as noted, there was some concern as to whether enough had been done to use the PSS data to improve SASS. Also, there was an uncomfortable degree of ad hoc tinkering.

If number of students was the major predictive variable, a sensible mass imputation method that could be applied would be to simply impute the SASS records to nearby PSS cases where nearness is defined simply by student enrollment. For parts of the distribution where the SASS sample is sparse, the SASS observation could be used over and over as a donor perhaps up to, say, 1.5 times its original SASS weight. Conversely, in parts of the distribution where there were lots of SASS cases relative to those in the PSS, the SASS cases would be used as donors less often than their original SASS weights would suggest, maybe only half as much.

The imputation or weight range, from about one half to about one and a half, is clearly arbitrary and depends on how much of a potential variance price one is willing to pay to get the "nearness" desired. In many weighting settings -- e.g., Oh and Scheuren (1987), however, these weight truncating factors seem to work well.

It may be useful to think of choosing a mass imputation approach after successively imputing SASS to each of the PSS variables separately. This way it would be possible to look at how often each SASS observation was used as a donor. If this range of donor use is not too large, then a single, perhaps nearest neighbor, imputation model might work well. Widely discrepant values in terms of donor use would suggest that the imputation is sensitive to one's beliefs as to the predictive power of the variables being used in the imputation. In such settings a case can be made for doing several different imputations that might be made available to the final users for possibly different uses.

4.4.4 What Next? -- The estimation problems tackled in this report are part of a larger set of issues for SASS as it evolves in a world where tight budgets may lead to less frequent large-scale surveys. SASS, for example, has already been shifted from a three-year to a five-year cycle. The groundwork for a complete rethinking of SASS has already been developed by numerous research efforts such as this (best summarized in a regular working paper series put together from papers given at professional meetings). The real priority is to put these efforts together into a larger redesign effort. While beyond the scope of the current report, it may be worth noting that some brainstorming has been done in this area (Scheuren, F., 1996a, 1996b). More is clearly needed. Most of the scenarios looked at so far, though, suggest that mass imputation be given serious study.

5. REFERENCES

Abramson, R., Cornette, C., Jackson, B., Palmer, R.and Kaufman, S.(1996), "!993-94 Schools and Staffing Survey: Sample Design and Estimation," *NCES Working Paper No. 96-089.*

Brewer, K. (1994), "Survey Sampling Inference: Some Past Perspectives and Present Prospects," *Pak. J. Statist.*, 10(1)A, 213-233.

Bankier, M. (1992), "Two-step Generalized Least Squares Estimation in the 1991 Canadian Census," in *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Broughman, S. (1996), 1993-94 Private School Universe Survey, *Statistical Analysis Report NCES 94-350* Washington, DC: U.S. Department of Education.

Broughman, S., Gerald, E., Bynum, L. and Stoner, K. (1994), Private School Universe Survey, 1991-92, National Center of Education Statistics, *Statistical Analysis Report NCES 94-350* Washington, DC: U.S. Department of Education.

Burton, R. (1989), Unpublished Memorandum, National Center for Education Statistics.

Causey, B., Bailey, L., and Hoy, E. (1996), "Alternative Methods of Coverage Estimation for the Private School Survey," Presented at the Joint Statistical Meetings in Chicago, August, 1996.

Casady, R. and Valliant, R.(1993), "Conditional Properties of Post-Stratified Estimators Under Normal Theory," *Survey Methodology,* 19, 183-192.

Cleveland, W. (1993), *Visualizing Data*, Hobart Press: Summit, New Jersey.

Clogg, C., Rubin, D., Schenker, N., Schultz, B., and Weidman, L. (1991), "Multiple Imputation of Industry and Occupation Codes in Census Public Use Samples using Bayesian Logistic Regression," *Journal of the American Statistical Association*, 86, 68-78.

Colledge, M., Johnson, J., Pare, R. and Sande, I. (1978), "Large Scale Imputation of Survey Data," *Proceedings of the Survey Research Methods Section, American Statistical Association.*

Cox, B. and Cohen, S. (1985). *Methodological Issues for Health Care Surveys*, Marcel Decker: New York.

Deming, W.E. and Stephan, F.F. (1940), "On a Least Squares Adjustment of a Sampled

Frequency When the Expected Marginal Tables are Known," *Annals of Mathematical Statistics,* 11, 427-444.

Deville, J.C., and Särndal, C.E. (1992), "Calibration Estimators in Survey Sampling," *Journal of the American Statistical Association,* 87, 376-382.

Deville, J.C., Särndal, C.E. and Sautory,O. (1993), "Generalized Raking Procedures in Survey Sampling," *Journal of the American Statistical Association,* 88, 1013-1020.

Folsom, R. (1981), "The Equivalence of Generalized Double Sampling Regression Estimators, Weight Adjustments and Randomized Hot Deck Imputations, " *Proceedings of the Survey Research Methods Section, American Statistical Association.*

Fuller, W., Loughin, M., and Baker, H. (1994), "Regression Weighting in the Presence of Nonresponsewith Application to the 1987-1988 Nationwide Food Consumption Survey," *Survey Methodology*, 20, 75-86.

Hansen, M., Hurwitz, W.,and Madow, W.(1953), *Sample Survey Methods and Theory*, New York: Wiley.

Hinkins S. and Scheuren, F. (1986), "Hot Deck Imputation Procedure Applied to a Double Sampling Scheme," *Survey Methodology*, 12, 181-96.

Holt, A., Kaufman, S., Scheuren, F. and Smith, W. (1994), "Intersurvey Consistency in School Surveys," *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Holt, D. and Smith, T.M.F. (1979), "Post-stratification," *Journal of the Royal Statistical Society, Series A,* 142, 33-46.

Huang, E. (1978), *Nonnegative Regression Estimation for Sample Survey Data,* Unpublished PhD Dissertation, Iowa State University.

Imbens, G.W. and Hellerstein, J.K. (1993), "Raking and Regression," *Discussion Paper Number 1658*, Cambridge, MA, Harvard Institute of Economic Research, Harvard University.

Ireland, C.T. and Kullback, S. (1968), "Contingency Tables with Known Marginals," *Biometrika,* 55, 179-188.

Jackson, B., Frazier, R., King, K., and Schwanz, D. (1994), Improving the Coverage of Private Elementary-Secondary Schools," *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Jayasuriya, B. and Valliant, R. (1995), "An Application of Regression and Calibration Estimation to Post-stratification in a Household Survey," *BLS Working Notes No. 38.* To appear in *Survey Methodology*, December 1996 issue.

Kaufman, S. (1995), "Properties of the Schools and Staffing Survey's Bootstrap Estimator," *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Kaufman, S. (1996), *Properties of the Schools and Staffing Survey's Bootstrap Estimator in Nearest Neighbor Matching.* Paper given at the Chicago meetings of the American Statistical Association.

Kaufman, S. and Huang, H. (1993), "1990-91 Schools and Staffing Survey: Sample Design and Estimation," *Technical Report NCES 91-127*, Washington, DC: U.S. Department of Education.

Kaufman, S., Li, B.,and Scheuren, F. (1995), "Improved GLS Estimation in NCES Surveys," *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Kaufman, S. And Scheuren, F. (1996), "Where Will It All End? GLS and Mass Imputation Compared," to appear in the *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Kott, P. (1996), Some Notes on Calibration Estimators Based on Several Separate Stratifications," (Unpublished).

Kovar, J. and Whitridge, P. (1995), "Imputation of Business Survey Data," in Cox, Binder, Chinnappa, Christianson, Colledge, and Kott, eds. *Business Survey Methods*, Wiley: New York.

Li, B. and Scheuren, F. (1995), "Intersurvey Consistency in NCES Private School Surveys," *NCES Working Paper No. 95-16.*

Little, R. (1991), "Post-Stratification: A Modeler's Perspective," *Proceedings of the Section on Survey Research Methods, American Statistical Association..*

Little, R. and Wu, M. (1991), "Models for Contingency Tables with Known Marginals When Target and Sample Populations Differ," *Journal of the American Statistical Association, American Statistical Association.*

McMillen, M. and Benson, P. (1991), Diversity of Private Schools, *Technical Report NCES 92-082*, Washington, DC: U.S. Department of Education.

Oh, H.L. and Scheuren, F. (1978a), "Multivariate Raking Ratio Estimation in the 1973 Exact Match Study," in *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Oh, H.L. and Scheuren, F. (1978b), "Some Unresolved Application Issues in Raking Ratio Estimation," in *Proceedings of the Section on Survey Research Methods, American*

Statistical Association.

Oh, H.L. and Scheuren, F. (1983), "Weighting Adjustments for Unit Nonresponse in Surveys," *Incomplete Data In Surveys*, Vol. 2, National Academy of Science.: Washington, DC.

Oh, H.L. and Scheuren, F. (1987), "Modified Raking Ratio Estimation in the Corporate Statistics of Income," *Survey Methodology, 13, 209-219.*

Olkin, I. (1958), "Multivariate Ratio Estimation for Finite Populations," *Biometrika*, 45,154-165.

Rubin, D. (1987), *Multiple Imputation*, New York: Wiley.

Rubin, D. (1996), "Multiple Imputation After 18+ Years," *Journal of the American Statistical Association.* June 1996 (with discussuion). See also Rao, J.N.K. (1996), "On Variance Estimation with Imputed Survey Data," *Journal of American Statistical Association* (a companion paper to Rubin's in the same issue).

Särndal, C.E., Swensson, B. and Wretman, J. (1992), *Model Assisted Survey Sampling,* New York: Springer-Verlag.

Scheuren, F. (1996a) "Administrative Record Opportunities in Education Survey Research," *From Data to Information: New Directions for the National Center for Education Statistics,* U.S. Department of Education: Washington, D.C. See also Scheuren, F. (1996b), "Improved Estimation in the Schools and Staffing Survey," Presented at the Joint Statistical Meetings in Chicago, August, 1996.

Scheuren, F., Monaco, D., Zhang, F., Ikosi G., and Chang, M. (1996), *An Exploratory Analysis of Response Rates, 1990-91 Schools and Staffing Survey (SASS).* Report prepared for U.S. Department of Education, National Center for Education Statistics.

Whitridge, P., Bureau, M., and Kovar, J. (1990), "Mass Imputation at Statistics Canada," *Proceedings of the Sixth Annual Research Conference,* Washington, D.C., U.S. Bureau of Census, 666-675.

Wong, W. and Ho, C. (1991), "Bootstrapping Post-Stratification and Regression Estimates from a Highly Skewed Distribution," *1991 Proceedings of the Section on Survey Research Methods, American Statistical Association.*

6. APPENDIX
ILLUSTRATIONS OF ALTERNATIVE APPROACHES

As noted, generalized least squares estimators can have many forms. This is true even within the specialized set of constraints that are to be imposed on the SASS. Three alternatives are illustrated here: The modified GLS that is one of the methods featured in the main body of this report, the Olkin version of the modified GLS (our current preference), and an alternative that initially offered promise, based on dividing SASS at the median for both students and separately for teachers. All of these are illustrated below.

To fix ideas, consider the following "toy" example that may help illustrate the differences between methods. First, suppose a SASS subgroup has ten observations; and even though this is probably too small, the methods discussed here are to be applied. Second, the observations appear below as column vectors where the components:

$$\begin{matrix} w_i \\ t_i \\ s_i \end{matrix}$$

correspond to schools, teachers, and students respectively. In particular, the SASS data are

1	1	1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8	9	10
1	6	2	7	3	8	4	9	5	10

Aggregating the three SASS components yields:

$$\begin{matrix} \Sigma w_i & & w. & & 10 \\ \Sigma t_i & = & t. & = & 55 \\ \Sigma s_i & & s. & & 55 \end{matrix}$$

Third, suppose the PSS totals for this subgroup are:

$$\begin{matrix} N & & 10 \\ T & = & 50 \\ S & & 50 \end{matrix}$$

Notice, the SASS school total has already been set equal to that in the PSS. This has been done so that the example starts where a standard SASS estimation procedure might end.

Modified GLS Method. -- For the "modified GLS" the elements of the matrix M and the vector $\underline{d}$ need to be obtained. It is immediate that $\underline{d}$ is:

10 - 10 = 0
50 - 55 = -5
50 - 55 = -5

For the matrix M, after some calculation, the values are:

$$\begin{matrix} 10 & 55 & 55 \\ 55 & 385 & 355 \\ 55 & 355 & 385 \end{matrix}$$

For the inverse of M^{-1}, the values turn out to be:

$$\begin{matrix} .5481 & -.0407 & -.0407 \\ -.0407 & .0204 & -.0130 \\ -.0407 & -.0130 & 0.0204 \end{matrix}$$

Thus, solving

$$\underline{\lambda} = M^{-1}\underline{d}$$

the vector is $\underline{\lambda}' = (.4074, -.0370, -.0370)$ and the modified GLS weights are of the form:

$$u_i = w_i + .4074 - .0370t_i - .0370s_i$$

Olkin GLS Method. -- The Olkin GLS Method is a slight variation on the modified GLS. Instead of just solving the matrix equation $\underline{\lambda} = M^{-1}\underline{d}$, as above, an overall weighted ratio adjustment r is made to the data first, such that the equality

$$r(a_1 R_1 w_{.} + a_2 R_2 t_{.} + a_3 R_3 s_{.}) = S$$

holds, where --

the lower case Roman letters w., t., and s. are the sample (SASS) estimates and

the upper case Roman letters N, T, and S are the target (PSS) values to be attained.

The R's are the target ratios

$$R_1 = S/N$$

$$R_2 = S/T$$

$$R_3 = S/S$$

and the a's are nonnegative and such that they add to one. In the present report, the a's have all been taken to be 1/3.

As discussed elsewhere, the "r" adjustment has the effect of making a weighted convex combination of the d_i's equal to zero. Intuitively, this was expected to reduce the number of negative weights; and, when done separately within subclasses, to achieve some of the usual benefits of post-stratification.

In the current illustration, the

$$R_1 = S/N = 50/10 = 5$$

$$R_2 = S/T = 50/50 = 1$$

$$R_3 = S/S = 50/50 = 1.$$

Hence,

$$r = 50/(1/3)(\ 50 + 55 + 55) = .9375.$$

This adjustment is then applied to the sample data before the matrix equation is solved.

Median GLS Method. -- To carry out this method, begin by dividing the SASS observations at the median value of the teachers t_i and then divide the SASS cases yet again at the median of the students s_i. Four groups are thus formed:

t_i, s_i both below median	t_i above median; s_i below
t_i below median; s_i above	t_i and s_i both above

An adjustment algorithm is developed by applying the intuitive idea that if SASS student estimates are, say, too small, then there are not enough large schools in the sample and thus those above the median should be reweighted up, by say $(1+\beta)$.

To keep the number of schools fixed, an equal but opposite adjustment $(1-\beta)$ is required for those schools below the median number of students. Similar considerations apply to an upward (downward) adjustment of $(1+\alpha)$ for SASS teacher estimates.

Unlike the modified GLS and Olkin GLS methods in the main report, this "median GLS" is iterative and requires repeated application of the adjustment process: first to the student totals, then to the teacher totals and so on. Each adjustment is to be made to the new cell totals derived from the previous adjustment(s).

To fix the specifics here, a detailed illustration is given using the same illustrative data as earlier:

1	1	1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8	9	10
1	6	2	7	3	8	4	9	5	10

Aggregating the three SASS components yields:

10
55
55

Now suppose the PSS totals for this subgroup are:

10
50
50

Notice, the SASS school total has already been set equal to that in the PSS. This has been done, as earlier noted, so that the example starts where a standard SASS estimation procedure might end.

In carrying out the "median GLS" method, the data are divided at the median for both teachers and schools. When this is done, the resulting data are arrayed as:

1	1
7	9
4	5

1	1	1
6	8	10
8	9	10

1	1	1
1	3	5
1	2	3

1	1
2	4
6	7

The corresponding cell totals are

2	3
16	24
9	27

3	2
9	6
6	13

To bring the second SASS component in line with the second PSS component an adjustment of the form:

2	3
$(1+\alpha)16$	$(1+\alpha)24$
9	27

3	3
$(1+\alpha)9$	$(1+\alpha)6$
6	13

is made. Solving for α,

$$\alpha = \frac{50-(16+24+9+6)}{(16+24)-(9+6)} = -1/5.$$

Substituting this value for alpha yields the following new cell totals

1.6	2.4
12.8	19.2
7.2	21.6

3.6	2.4
10.8	7.2
7.2	15.6

and the corresponding overall totals have become

10.0
50.0
51.1

To bring the third (student) SASS component in line with the hypothetical PSS total, the adjustment proceeds this time by columns where:

1.6	2.4
$(1+\beta)$ 12.8	$(1-\beta)$ 19.2
7.2	21.6

3.6	2.4
$(1+\beta)$ 10.8	$(1-\beta)$ 7.2
7.2	15.6

Solving for β the expression obtained is

$$\beta = \frac{50-(21.6+15.6+7.2+7.2)}{(21.6+15.6)-(7.2+7.2)}$$

$$= -1.5/22.8 \quad = -0.066.$$

After the adjustment, the new overall totals have become

10.2
49.8
50.0

The school totals are slightly out of balance and themselves may need adjustment; notice, too, that the teacher totals are off a bit but the gap is still smaller than the gap for students that was just removed. Continuing to cycle here would eventually yield SASS estimates that agreed to whatever closeness was desired with their corresponding PSS counterparts.

What then are the impacts on the weights, assuming the iteration was stopped at this point? The adjustments are

$$(1+\alpha)(1+\beta) = (0.800)(1.066) = 0.853$$

$$(1+\alpha)(1-\beta) = (0.800)(0.934) = 0.750$$

$$(1-\alpha)(1+\beta) = (1.200)(1.066) = 1.280$$

$$(1-\alpha)(1-\beta) = (1.200)(0.934) = 1.120$$

As noted earlier, one measure of the weight variation caused by imposing these constraints is to calculate the average sum of squared weights. In this case, that sum turns out to be approximately

Modified GLS = 10.38

Median GLS = 10.56

or not a lot greater than the sum of the unadjusted squared weights (at 10.00). In other words, there is not much to choose from here between the two methods.

Further Considerations.-- The median GLS method just described was tried on two of the typologies in the SASS private school component (Li and Scheuren 1995). For the first of these, the Catholic Private component, the technique worked reasonably satisfactorily but for the second typology, the Non-sectarian Special Emphasis component, the algorithm did not converge. Apparently when negative weights arise in the modified GLS (used in the main body of this report) the median GLS may not converge due to the inconsistency. Because of this experience the approach was abandoned.

Listing of NCES Working Papers to Date

Please contact Ruth R. Harris at (202) 219-1831
if you are interested in any of the following papers

Number	Title	Contact
94-01 (July)	Schools and Staffing Survey (SASS) Papers Presented at Meetings of the American Statistical Association	Dan Kasprzyk
94-02 (July)	Generalized Variance Estimate for Schools and Staffing Survey (SASS)	Dan Kasprzyk
94-03 (July)	1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report	Dan Kasprzyk
94-04 (July)	The Accuracy of Teachers' Self-reports on their Postsecondary Education: Teacher Transcript Study, Schools and Staffing Survey	Dan Kasprzyk
94-05 (July)	Cost-of-Education Differentials Across the States	William Fowler
94-06 (July)	Six Papers on Teachers from the 1990-91 Schools and Staffing Survey and Other Related Surveys	Dan Kasprzyk
94-07 (Nov.)	Data Comparability and Public Policy: New Interest in Public Library Data Papers Presented at Meetings of the American Statistical Association	Carrol Kindel
95-01 (Jan.)	Schools and Staffing Survey: 1994 Papers Presented at the 1994 Meeting of the American Statistical Association	Dan Kasprzyk
95-02 (Jan.)	QED Estimates of the 1990-91 Schools and Staffing Survey: Deriving and Comparing QED School Estimates with CCD Estimates	Dan Kasprzyk
95-03 (Jan.)	Schools and Staffing Survey: 1990-91 SASS Cross-Questionnaire Analysis	Dan Kasprzyk
95-04 (Jan.)	National Education Longitudinal Study of 1988: Second Follow-up Questionnaire Content Areas and Research Issues	Jeffrey Owings
95-05 (Jan.)	National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors	Jeffrey Owings

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
95-06 (Jan.)	National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data	Jeffrey Owings
95-07 (Jan.)	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
95-08 (Feb.)	CCD Adjustment to the 1990-91 SASS: A Comparison of Estimates	Dan Kasprzyk
95-09 (Feb.)	The Results of the 1993 Teacher List Validation Study (TLVS)	Dan Kasprzyk
95-10 (Feb.)	The Results of the 1991-92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation	Dan Kasprzyk
95-11 (Mar.)	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
95-12 (Mar.)	Rural Education Data User's Guide	Samuel Peng
95-13 (Mar.)	Assessing Students with Disabilities and Limited English Proficiency	James Houser
95-14 (Mar.)	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
95-15 (Apr.)	Classroom Instructional Processes: A Review of Existing Measurement Approaches and Their Applicability for the Teacher Follow-up Survey	Sharon Bobbitt
95-16 (Apr.)	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95-17 (May)	Estimates of Expenditures for Private K-12 Schools	Stephen Broughman
95-18 (Nov.)	An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey	Dan Kasprzyk
96-01 (Jan.)	Methodological Issues in the Study of Teachers' Careers: Critical Features of a Truly Longitudinal Study	Dan Kasprzyk

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-02 (Feb.)	Schools and Staffing Survey (SASS): 1995 Selected papers presented at the 1995 Meeting of the American Statistical Association	Dan Kasprzyk
96-03 (Feb.)	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
96-04 (Feb.)	Census Mapping Project/School District Data Book	Tai Phan
96-05 (Feb.)	Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey	Dan Kasprzyk
96-06 (Mar.)	The Schools and Staffing Survey (SASS) for 1998-99: Design Recommendations to Inform Broad Education Policy	Dan Kasprzyk
96-07 (Mar.)	Should SASS Measure Instructional Processes and Teacher Effectiveness?	Dan Kasprzyk
96-08 (Apr.)	How Accurate are Teacher Judgments of Students' Academic Performance?	Jerry West
96-09 (Apr.)	Making Data Relevant for Policy Discussions: Redesigning the School Administrator Questionnaire for the 1998-99 SASS	Dan Kasprzyk
96-10 (Apr.)	1998-99 Schools and Staffing Survey: Issues Related to Survey Depth	Dan Kasprzyk
96-11 (June)	Towards an Organizational Database on America's Schools: A Proposal for the Future of SASS, with comments on School Reform, Governance, and Finance	Dan Kasprzyk
96-12 (June)	Predictors of Retention, Transfer, and Attrition of Special and General Education Teachers: Data from the 1989 Teacher Followup Survey	Dan Kasprzyk
96-13 (June)	Estimation of Response Bias in the NHES:95 Adult Education Survey	Steven Kaufman
96-14 (June)	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman
96-15 (June)	Nested Structures: District-Level Data in the Schools and Staffing Survey	Dan Kasprzyk

BEST COPY AVAILABLE

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-16 (June)	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
96-17 (July)	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio
96-18 (Aug.)	Assessment of Social Competence, Adaptive Behaviors, and Approaches to Learning with Young Children	Jerry West
96-19 (Oct.)	Assessment and Analysis of School-Level Expenditures	William Fowler
96-20 (Oct.)	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96-21 (Oct.)	1993 National Household Education Survey (NHES:93) Questionnaires: Screener, School Readiness, and School Safety and Discipline	Kathryn Chandler
96-22 (Oct.)	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
96-23 (Oct.)	Linking Student Data to SASS: Why, When, How	Dan Kasprzyk
96-24 (Oct.)	National Assessments of Teacher Quality	Dan Kasprzyk
96-25 (Oct.)	Measures of Inservice Professional Development: Suggested Items for the 1998-1999 Schools and Staffing Survey	Dan Kasprzyk
96-26 (Nov.)	Improving the Coverage of Private Elementary-Secondary Schools	Steven Kaufman
96-27 (Nov.)	Intersurvey Consistency in NCES Private School Surveys for 1993-94	Steven Kaufman

DEPARTMENT OF EDUCATION
UNITED STATES OF AMERICA